Guaranteed Matrix Completion via Non-Convex Factorization

Ruoyu Sun and Zhi-Quan Luo, *Fellow, IEEE*

***Abstract*— Matrix factorization is a popular approach for large-scale matrix completion. The optimization formulation based on matrix factorization, even with huge size, can be solved very efficiently through the standard optimization algorithms in practice. However, due to the non-convexity caused by the factorization model, there is a limited theoretical understanding of whether these algorithms will generate a good solution. In this paper, we establish a theoretical guarantee for the factorization-based formulation to correctly recover the underlying low-rank matrix. In particular, we show that under similar conditions to those in previous works, many standard optimization algorithms converge to the global optima of a factorization-based formulation and recover the true low-rank matrix. We study the local geometry of a properly regularized objective and prove that any stationary point in a certain local region is globally optimal. A major difference of this paper from the existing results is that we do not need resampling (i.e., using independent samples at each iteration) in either the algorithm or its analysis.**

***Index Terms*— Matrix completion, matrix factorization, nonconvex optimization, alternating minimization, SGD, perturbation analysis.**

I. Introduction

In THE era of big data, there has been an increasing need for handling the enormous amount of data generated by mobile devices, sensors, online merchants, social networks, etc. Exploiting low-rank structure of the data matrix is a powerful method to deal with "big data". One prototype example is the low rank matrix completion problem in which the goal is to recover an unknown low rank matrix $M \in \mathbb{R}^{m\times n}$ for which only a subset of its entries $M_{ij}, (i, j) \in \Omega \subseteq \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$ are specified. Matrix completion has found numerous applications in various fields such as recommender systems [1], computer vision [2] and system identification [3], to name a few.

There are two popular approaches to impose the low-rank structure: the nuclear norm based approach and the matrix factorization (MF) based approach. In the first approach, the whole matrix is the optimization variable and the nuclear norm (denoted as $\|\cdot\|_*$) of this matrix variable, which can be viewed as a convex approximation of its rank, serves as the objective function or a regularization term. For the matrix completion problem, the nuclear norm based formulation becomes either a linearly constrained minimization problem [4]

$$\min_{Z\in\mathbb{R}^{m\times n}} \|Z\|_*, \quad \text{s.t. } Z_{ij} = M_{ij}, \quad \forall (i, j) \in \Omega, \tag{1}$$

a quadratically constrained minimization problem

$$\min_{Z\in\mathbb{R}^{m\times n}} \|Z\|_*, \quad \text{s.t.} \sum_{(i,j)\in\Omega} (Z_{ij} - M_{ij})^2 \le \epsilon, \tag{2}$$

or a regularized unconstrained problem

$$\min_{Z\in\mathbb{R}^{m\times n}} \|Z\|_* + \lambda \sum_{(i,j)\in\Omega} (Z_{ij} - M_{ij})^2. \tag{3}$$

On the theoretical side, it has been shown that given a rank-r matrix M satisfying an incoherence condition, solving (1) will exactly reconstruct M with high probability provided that $O(r(m+n)\log^2(m+n))$ entries are uniformly randomly revealed [4]–[7]. This result was later generalized to noisy matrix completion, whereby the optimization formulation (2) is adopted [8]. Using a different proof framework, reference [9] provided theoretical guarantee for a variant of the formulation (3). On the computational side, problems (1) and (2) can be reformulated as a semidefinite program (SDP) and solved to global optima by standard SDP solvers when the matrix dimension is smaller than 500. To solve problems with larger size, researchers have developed first order algorithms, including the SVT (singular value thresholding) algorithm for the formulation (1) [10], and several variants of the proximal gradient method for the formulation (3) [11], [12]. Although linear convergence of the proximal gradient method has been established for the formulation (3) under certain conditions [13], [14], the per-iteration cost of computing SVD (Singular Value Decomposition) may increase rapidly as the dimension of the problem increases, making these algorithms rather slow or even useless for problems of huge size. The other major drawback is the memory requirement of storing a large m by n matrix.

In the second approach, the unknown rank r matrix is expressed as the product of two much smaller matrices XY^T, where $X \in \mathbb{R}^{m\times r}$, $Y \in \mathbb{R}^{n\times r}$, so that the low-rank requirement is automatically fulfilled. Such a matrix factorization model has long been used in PCA (principle component analysis)

Manuscript received January 8, 2015; revised September 29, 2015; accepted May 29, 2016. Date of publication August 8, 2016; date of current version October 18, 2016. This work was supported in part by NSF under Grant CCF-1526434, in part by NSFC under Grant 61571384, and in part by a Doctoral Dissertation Fellowship from the Graduate School of the University of Minnesota. This paper was presented in part at the 2015 IEEE FOCS.

R. Sun is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: sunxx394@umn.edu).

Z.-Q. Luo is with The Chinese University of Hong Kong, Hong Kong, and also with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: luozq@cuhk.edu.cn).

Communicated by V. Saligrama, Associate Editor for Signal Processing.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2016.2598574

and many other applications [15]. It has gained great popularity in the recommender systems field and served as the basic building block of many competing algorithms for the Netflix Prize [1], [16] due to several reasons. First, the compact representation of the unknown matrix greatly reduces the per-iteration computation cost as well as the storage space (requiring essentially linear storage of $O((m+n)r)$ for small r). Second, the per-iteration computation cost is rather small and people have found in practice that huge size optimization problems based on the factorization model can be solved very fast. Third, as elaborated in [1], the factorization model can be easily modified to incorporate additional application-specific requirements.

A popular factorization based formulation for matrix completion takes the form of an unconstrained regularized square-loss minimization problem [1]:

$$\begin{aligned} \text{P0}: \min_{X\in\mathbb{R}^{m\times r}, Y\in\mathbb{R}^{n\times r}} & \frac{1}{2}\sum_{(i,j)\in\Omega}[M_{ij}-(XY^T)_{ij}]^2 \\ & + \lambda(\|X\|_F^2+\|Y\|_F^2). \end{aligned} \tag{4}$$

There are a few variants of this formulation: the coefficient λ can be zero [17]–[20] or different for each row of X, Y [21]; each square loss term $[M_{ij}-(XY^T)_{ij}]^2$ can have different weights [1]; an additional matrix variable $Z \in \mathbb{R}^{n\times r}$ can be introduced [22]. Problem (4) is a non-convex fourth-order polynomial optimization problem, and can be solved to stationary points by standard nonlinear optimization algorithms such as gradient descent method, alternating minimization [1], [18], [19], [21] and SGD (stochastic gradient descent) [1], [16], [23], [24]. Alternating minimization is easily parallelizable but has higher per-iteration computation cost than SGD; in contrast, SGD requires little computation per iteration, but its parallelization is challenging. Recently several parallelizable variants of the SGD [25]–[27] and variants of the block coordinate descent method with very low per-iteration cost [28], [29] have been developed. Some of these algorithms have been tested in distributed computation platforms and can achieve good performance and high efficiency, solving very large problems with more than a million rows and columns in just a few minutes.

A. Our Contributions

Despite the great empirical success, the theoretical understanding of the algorithms for the factorization based formulation is fairly limited. More specifically, the fundamental question of whether these algorithms (including many recently proposed ones) can recover the true low-rank matrix remains largely open. In this paper, we partially answer this question by showing that under similar conditions to those used in previous works, many standard optimization algorithms for a factorization based formulation (see (18)) indeed converge to the true low-rank matrix (see Theorem 3.1). Our result applies to a large class of algorithms including gradient descent, SGD and many block coordinate descent type methods such as two-block alternating minimization and block coordinate gradient descent. We also show the linear convergence of some of these algorithms (see Theorem 3.2 and Corollary 3.2).

To the best of our knowledge, our result is the first one that analyzes the geometry of matrix factorization in Euclidean space for matrix completion. In addition, our result also provides the first recovery guarantee for alternating minimization without resampling (i.e. without using independent samples in different iterations). Below we elaborate these two contributions in light of the existing works.

1) We analyze the local geometry of the matrix factorization formation (in Euclidean space). Our result provides a validation of the formulation rather than a validation of a single algorithm. In other words, the success of many algorithms attributes mostly to the geometry of the problem, rather than the specific algorithms being used. We analyze the local geometry of a classical formulation $\|M - XY^T\|_F^2$ by a novel perturbation analysis, then for the sampling loss $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2$ we show how to add regularizers to establish a similar local geometrical property. A high-level lesson is that regularization may change the geometry of the problem.

2) Our result applies to the standard forms of the algorithms (though our optimization formulation is a bit different), which do not require the additional resampling scheme used in other works [17]–[20]. We obtain a sample complexity bound that is independent of the recovery error ϵ, while all previous sample complexity bounds for the matrix factorization based formulation (in Euclidean space) depend on ϵ. There is a subtle theoretical issue for the resampling scheme; see more discussions in Section I-B and [30, Sec. 1.5.3].

B. Related Works

1) Factorization Models: The first recovery guarantee for the factorization based matrix completion is provided in [31], where Keshavan, Montanari and Oh considered a factorization model in Grassmannian manifold and showed that the matrix can be recovered by a proper initialization and a gradient descent method on Grassmannian manifold. Besides being quite complicated, this model is not as flexible as the factorization model in Euclidean space, and it is not easy to solve by many advanced large-scale optimization algorithms. Moreover, most algorithms in Grassmann manifold require line search, and little is known about the convergence rate.

The factorization model in Euclidean space was first analyzed in an unpublished work [17] of Keshavan,[1] as well as a later work of Jain et al. [18]. Both works considered alternating minimization with resampling scheme, a special variant of the original alternating minimization. The sample complexity bounds were later improved by Hardt [19] and Hardt and Wootters [20], where in the latter work, notably, the authors devised an algorithm with a corresponding sample complexity bound independent of the condition number. However, these improvements are obtained for more sophisticated versions of resampling-based alternating minimization, not the typical alternating minimization algorithm.

[1]Reference [17] is a PhD thesis that discusses various algorithms including the algorithm proposed in [31] and alternating minimization. In this paper when we refer to [17], we are only referring to [17, Ch. 5] which presents resampling-based alternating minimization and the corresponding result.

2) Resampling: The issues of resampling have been discussed in a recent work on phase retrieval by Candès et al. [32]. We will point out a subtle theoretical issue not mentioned in [32], as well as some other practical issues.

The resampling scheme (a.k.a. golfing scheme [6]) can be used at almost no cost for the nuclear norm approach [6], [7], [33], but for the alternating minimization it causes many issues. At first, it may seem that for both approaches resampling is a cheap way to get around a common difficulty: the dependency of the iterates on the sample set. However, there is a crucial difference: for the nuclear norm approach, resampling is just a proof technique used in a "conceptual" algorithm for constructing the dual certificate, while for the alternating minimization, resampling is used in the actual algorithm. This difference causes some issues of resampling-based alternating minimization at conceptual, practical and theoretical levels.

1) Gap between theory and algorithm. Algorithmically, an easy resampling scheme is to randomly partition the given set Ω into non-overlapping subsets $\Omega_k, k = 1, \ldots, L$, as proposed in [17] and [18].[2] However, the results in [17]–[20] actually require a generative model of independent Ω_k's, instead of sampling Ω_k's based on a given Ω. Therefore, the results in [17]–[20] do not directly apply to the partition based resampling scheme that is easy to use. See [30, Sec. 1.5.3] for more discussions on this subtle issue.

This issue has been discussed by Hardt and Wooters in [20, Appendix D], and they proposed a new resampling scheme [20, Algorithm 6] to which the results in [17]–[20] can apply, provided that the generative model of Ω is exactly known. In practice, the underlying generative model of Ω is usually unknown, in which case the scheme [20, Algorithm 6] does not work. In contrast, the classical results in [4]–[7] and our result herein are robust to the generative model of Ω: these results actually state that for an overwhelming portion of Ω with a given size, one can recover M through a certain algorithm, thus for many reasonable probability distributions of Ω a high probability result holds.

2) Impracticality. As argued previously, assuming a generative model of Ω_k's is not practical since Ω is usually given. For given Ω, the only known validated resampling scheme [20, Algorithm 6], besides not being robust to the underlying generative model of Ω, might be a bit complicated to use in practice. Even the simple resampling scheme of partitioning Ω (which has not been validated yet) is rather unrealistic since each sample is used only once during the algorithm.

3) Inexact recovery. A theoretical consequence of the resampling scheme is that the required sample complexity $|\Omega|$ becomes dependent on the desired accuracy ϵ, and goes to infinity as ϵ goes to zero. This is different from the classical results (and ours) where exact reconstruction only requires finite samples. While it is common to see the dependency of *time complexity* on the accuracy ϵ, it is relatively uncommon to see the dependency of *sample complexity* on ϵ.

[2]The description in [18] has some ambiguity and it might refer to the scheme of sampling Ω_k's with replacement; anyhow, under this model Ω_k's are still dependent. See [30, Sec. 1.5.3] for more discussions.

In a recent work [34] the authors have managed to remove the dependency of the required sample size on ϵ by using a singular value projection algorithm. However, [34] considers a matrix variable of the same size as the original matrix, which requires significantly more memory than the matrix factorization approach considered in this paper. Moreover, it requires resampling at a number of iterations (though not all), which may suffer from the same issues we mentioned earlier. The resampling is also required in the recent work of [35]; see [30, Sec. 1.5.3] for more discussions.

3) Other Works on Non-Convex Formulations: Non-convex formulation has also been studied for the phase retrieval problem in some recent works [32], [36]. These works provide theoretical guarantee for some algorithms specially tailored to certain non-convex formulations and with specific initializations. The major difference between [32] and [36] is that the former requires independent samples in each iteration, while the latter uses the same samples throughout in the proposed algorithm. As mentioned earlier, such a difference also exists between all previous works on alternating minimization for matrix completion [17]–[20] and our work.

Finally, we note that there is a growing list of works on the theoretical guarantee of non-convex formulations for various problems, such as sparse regression (e.g. [37]–[39]), sparse PCA [40], [41], robust PCA [42] and EM (Expected-Maximization) algorithm [43], [44]. The techniques used in these works, however, seem to be quite different from those in the current work.

C. Proof Overview and Techniques

1) Basic Idea: Local Geometry: The very first question is what kind of property can ensure global convergence for non-convex optimization. We will establish a local geometrical property of a regularized objective such that any stationary point in a local region is globally optimal. This is achieved in three steps: (i) study the local geometry of the fully observed objective $\|M - XY^T\|_F^2$; (ii) study the local geometry of the matrix completion objective $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2$; (iii) study the local geometry of a regularized objective. Next, we will discuss the difficulties involved in each step and describe how we address these difficulties.

2) Local Geometry of $\|M - XY^T\|_F^2$: We start by considering a simple case that M is fully observed and the objective function is $f(X, Y) = \|M - XY^T\|_F^2$. What is the geometrical landscape of this function? In the simplest case $m = n = r = 1$ and $f(x, y) = (xy - 1)^2$, the set of stationary points is $\{(x, y) \mid xy = 1\} \cup \{(0, 0)\}$, in which $(0, 0)$ is a saddle point and the curve $xy = 1$ consists of global optima. We plot the function around the curve $xy = 1$ in the positive orthant in Figure 1.

Clearly a certain geometrical property prevents bad local minima in the neighborhood of the global optima, but what kind of property? We emphasize that the property can *not* be local convexity because the set of global optima is non-convex in $\mathbb{R}^2$. Due to the intrinsic symmetry that $f(x, y) = f(xq, yq^{-1})$, only the product $z = xy$ affects the value of f. We hope that the strong convexity of $(1 - z)^2$ can be partially

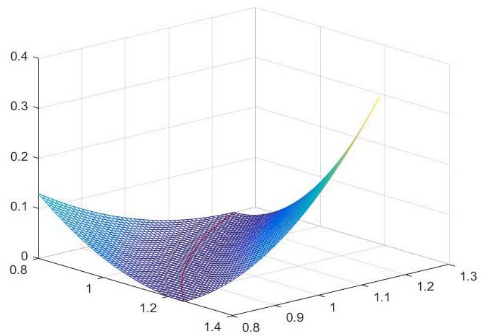

Fig. 1. The plot of function $f(x, y) = (xy - 1)^2$ around the set of global optima $xy = 1$ in the positive orthant. The bottom of this bowl shape is a hyperbola $xy = 1$.

preserved when z is reparameterized into $z = xy$. It turns out we can prove the following local convexity-type property: for any (x, y) such that xy is close to 1 and $|x|, |y|$ are upper bounded, there exists $uv = 1$ such that

$$\langle \nabla f(x, y), (x, y) - (u, v)\rangle \geq c\|(x, y) - (u, v)\|^2.$$

An interpretation is that the negative gradient direction $-\nabla f$ should be aligned with the global direction $(u, v) - (x, y)$; a convex function has a similar property, but the difference is that here the global direction is adjusted according to the position of (x, y).

For general m, n, r, the geometrical landscape is probably much more complicated than the scalar case. Nevertheless, we can still prove that the convexity of $\|M - Z\|^2$ is partially preserved when reparameterizing Z as $Z = XY$. The exact expression is a variant of (6) which we will discuss in more detail later. Technically, we need to connect the Euclidean space and the quotient manifold via "coupled perturbation analysis": given X, Y such that $\|XY^T - M\|_F$ is small, find decomposition $M = UV^T$ such that U, V are close to X and Y respectively (a simpler version of Proposition 4.1). The difference from traditional perturbation analysis of Wedin [45] (i.e. if two matrices are close then their row/column spaces are close) is that in [45] the row/column spaces are fixed while in our problem U, V are up to our choice.

3) Local Geometry of $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2$: Let us come back to the original matrix completion problem, in which an additional sampling operator $\mathcal{P}_\Omega$ is introduced. Similarly, we hope that $f_\Omega(Z) = \frac{1}{2}\|\mathcal{P}_\Omega(M - Z)\|^2$ is strongly convex and this property can be partially preserved after reparametrization $Z = XY^T$. However, one issue is that the function $f_\Omega(Z)$ is possibly non-strongly-convex (though still convex). In fact, if f_Ω is locally strongly convex around M, then we should have

$$f_\Omega(Z) - f_\Omega(M) \geq O(\|Z - M\|_F^2), \quad \forall Z \text{ close to } M.$$

Assuming Z is rank-r, this inequality can be rewritten as

$$\|\mathcal{P}_\Omega(M - XY^T)\|_F^2 \geq Cp\|M - XY^T\|_F^2, \quad \forall (X, Y) \in K(\delta), \tag{5}$$

where $K(\delta)$ is a neighborhood of M defined as $\{(X, Y) \mid \|XY^T - M\|_F \leq \delta\}$ and C is a numerical constant. We wish (5) to hold with high probability (w.h.p.) for random Ω in which each position in M is chosen with probability p. This inequality is closely related to matrix RIP (restricted isometry property) in [8] (see equation (III.4) therein). If X, Y are independent of Ω, then (5) follows easily from the concentration inequalities. Unfortunately, if X, Y are chosen arbitrarily instead of independently from Ω, the bound (5) may fail to hold.

A solution, as employed in [31], is to utilize a random graph lemma in [46] which provides a bound on $\|\mathcal{P}_\Omega(A)\|_F$ for any rank-1 matrix A (possibly dependent on Ω). This lemma, combined with another probability result in [4], implies a bound on $\|\mathcal{P}_\Omega(M - XY^T)\|_F$. However, this bound is not good enough since it only leads to (5) when $\delta = O(1/n)$. The underlying reason is that the bound given by the random graph lemma is actually quite loose if X or Y have unbalanced rows, i.e. certain row has large norm. One solution is to force the iterates to have bounded row norms (a.k.a. incoherent), by adding a constraint or regularizer. With the incoherence requirement on X, Y, now (5) can be shown to be hold for $\delta = O(1)$, or more precisely, $\delta = O(\Sigma_{\min})$, where $\Sigma_{\min}$ is the minimum eigenvalue of M. With such a δ, it is possible to find an initial point in the region $K(\delta)$.

In summary, although $f_\Omega(Z) = \frac{1}{2}\|\mathcal{P}_\Omega(Z - M)\|_F^2$ is possibly non-strongly-convex, by restricting to an incoherent neighborhood of M it is "relative" strongly convex (called "relative" since we fix M in (5)). More specifically, we have that w.h.p.

$$\begin{aligned}\|\mathcal{P}_\Omega(M - XY^T)\|_F^2 &\geq Cp\|M - XY^T\|_F^2,\\ \forall (X, Y) &\in \mathcal{B} \triangleq K(\delta) \cap K_1.\end{aligned} \tag{6}$$

where K_1 denotes the set of (X, Y) with bounded row norms. Note that this inequality also implies that global optimally in $\mathcal{B}$ leads to exact recovery; or equivalently, zero training error leads to zero generalization error.

Having established the geometry of $f_\Omega(Z)$, we can use the same technique for the fully observed case to show the local geometry[3] of

$$F(X, Y) \triangleq f_\Omega(X, Y) = \frac{1}{2}\|\mathcal{P}_\Omega(M - XY^T)\|_F^2.$$

More specifically, we can prove that for any $(X, Y) \in \mathcal{B}$, there exists $(U, V) \in \mathcal{X}^* = \{(U, V) \in \mathbb{R}^{m\times r} \times \mathbb{R}^{n\times r} \mid UV^T = M\}$ such that

$$\begin{aligned}&\langle \nabla_X F(X, Y), X - U\rangle + \langle \nabla_Y F(X, Y), Y - V\rangle\\ &\quad\geq c(\|X - U\|_F^2 + \|Y - V\|_F^2).\end{aligned} \tag{7}$$

Denoting $\boldsymbol{x} = (X, Y), \boldsymbol{x}^* = (U, V)$ and utilizing $\nabla F(\boldsymbol{x}^*) = 0$, (7) becomes

$$\begin{aligned}&\forall \boldsymbol{x} \in \mathcal{B}, \exists \boldsymbol{x}^* \in \mathcal{X}^*,\\ &\quad \text{s.t. } \langle \nabla F(\boldsymbol{x}) - \nabla F(\boldsymbol{x}^*), \boldsymbol{x} - \boldsymbol{x}^*\rangle \geq c\|\boldsymbol{x} - \boldsymbol{x}^*\|^2.\end{aligned} \tag{8}$$

[3]For illustration purpose, we present a two-step approach: first establish a geometrical property of $f_\Omega(Z)$, then extend the property to $f_\Omega(XY^T)$. However, our current proof does not follow the two-step approach but directly establish the property of $f_\Omega(XY^T)$. In fact, although we establish the property of $f_\Omega(Z)$ in Claim 3.1, the proof of this claim is very similar to the proof of (7).

It links the local optimality measure $\|\nabla F(\boldsymbol{x})\|$ with the global optimality measure $\text{dist}(\boldsymbol{x}, \mathcal{X}^*) = \min_{\boldsymbol{x}^* \in \mathcal{X}^*} \|\boldsymbol{x} - \boldsymbol{x}^*\|$, and implies that any stationary point of F in $\mathcal{B}$ is a global minimum.

If (8) holds for arbitrary $\boldsymbol{x}, \boldsymbol{x}^*$ then F would be strongly convex in $\boldsymbol{x}$. Let us emphasize again two differences of (8) with local strong convexity: i) since $\boldsymbol{x}^*$ is not arbitrary but has to be one global minimum, (8) indicates local "relative convexity" of F; ii) due to the ambiguity of factorization, $\boldsymbol{x}^*$ should be chosen according to $\boldsymbol{x}$, thus (8) indicates local relative convexity up to a group transformation (it might be conceptually helpful to view it as a property in the quotient manifold, but we do not explicitly exploit its structure).

4) Local Geometry With Regularizers/Constraints: The property (8) is still not desirable. The original purpose of studying geometry is to show there is no spurious "1st order local-min" (point that satisfies 1st order optimality conditions). To establish the geometrical property with sampling, we restrict to an incoherent set K_1, but this restriction changes the meaning of the 1st order local-min. In fact, to ensure the iterates stay in the incoherent region K_1, we need to solve a constrained optimization problem $\min_{\boldsymbol{x} \in K_1} F(\boldsymbol{x})$ or a regularized problem $\min_{\boldsymbol{x}} F(\boldsymbol{x}) + G_1(\boldsymbol{x})$ where G_1 is a regularizer forcing $\boldsymbol{x}$ to be in K_1. Standard optimization algorithms converge to the KKT points of $\min_{\boldsymbol{x} \in K_1} F(\boldsymbol{x})$ or the stationary points of $F + G_1$, which may not be the stationary points of F. The property (8) only implies any stationary point of F in $\mathcal{B}$ is globally optimal.

We shall focus on the regularized problem $\min F + G_1$; the constrained problem $\min_{\mathbf{x} \in K_1} F$ is similar. Because of the extra regularizer, the property (8) is not enough. We need to prove a result similar to (8), but with ∇F replaced by $\nabla F + \nabla G_1$:

$$\forall \boldsymbol{x} \in \mathcal{B},\ \exists\, \boldsymbol{x}^* \in \mathcal{X}^*,\quad \text{s.t. } \langle \nabla F(\boldsymbol{x}) + \nabla G_1(\boldsymbol{x}), \boldsymbol{x} - \boldsymbol{x}^* \rangle \geq c\|\boldsymbol{x} - \boldsymbol{x}^*\|^2. \tag{9}$$

If it happens to be the case that

$$\langle \nabla G_1(\boldsymbol{x}), \boldsymbol{x} - \boldsymbol{x}^* \rangle \geq 0, \tag{10}$$

then combining with the existing result (8) we are done; unfortunately, we do not know how to prove (10). Intuitively, (10) means that $-\nabla G_1(\boldsymbol{x})$, which is almost the same direction as the projection to the incoherent region K_1, is positively correlated with the global direction $\boldsymbol{x}^* - \boldsymbol{x}$. At first sight, this seems trivially true because for any point $\bar{\boldsymbol{x}} \in K_1$ we have $\langle \nabla G_1(\boldsymbol{x}), \boldsymbol{x} - \bar{\boldsymbol{x}} \rangle \geq 0$ (as illutrated in Fig. 2). However, a rather strange issue is that $\boldsymbol{x}^*$ is chosen to be a point in $\{(U, V) \mid UV^T = M\}$ that is close to $\boldsymbol{x}$, thus there is no guarantee that $\boldsymbol{x}^*$ lies in K_1. An underlying reason is that the global optimum set $\{(U, V) \mid UV^T = M\}$ is unbounded and thus not a subset of K_1. If we enforce (U, V) to be in K_1, we may not be able to find (U, V) that is close enough to (X, Y).

Technically, the issue is that (U, V) chosen in Proposition 4.1 have row-norms bounded above by quantities proportional to the norms of X, Y, and can be higher than the row-norms of X, Y (threshold of K_1). To resolve this issue, we add an extra regularizer $G_2(X, Y)$ to force (X, Y) to lie in K_2, a set of matrix pairs with bounded norms. This extra bound makes $\langle \nabla G_1(\boldsymbol{x}), \boldsymbol{x} - \boldsymbol{x}^* \rangle \geq 0$ straightforward to prove, but a similar issue arises: now we need to prove (8) for $F + G_1 + G_2$ instead of F. Again, it suffices to prove that for any $\boldsymbol{x} \in K(\delta) \cap K_1 \cap K_2$ there exists $\boldsymbol{x}^*$ such that

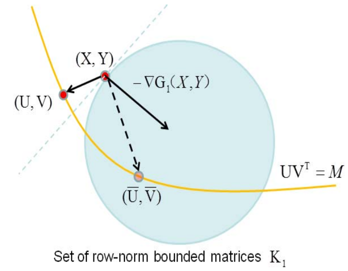

Fig. 2. Illustration of why a single regularizer G_1 is not enough. The requirement (10) means $-\nabla G_1(\boldsymbol{x}) = -\nabla G_1(X, Y)$ is positively correlated with $\boldsymbol{x}^* - \boldsymbol{x}$. This holds if we could pick some $\boldsymbol{x}^* = (\bar{U}, \bar{V})$ lying in the row-bounded region K_1. However, we need to choose $\boldsymbol{x}^* = (U, V)$ in the hyperbolic space $\{(U, V) \mid UV^T = M\}$ that is close to (X, Y). The figure indicates that such a (U, V) may be outside of K_1, and $(U, V) - (X, Y)$ may be negatively correlated with $-\nabla G_1(X, Y)$.

$$\langle \nabla G_2(\boldsymbol{x}), \boldsymbol{x} - \boldsymbol{x}^* \rangle \geq 0. \tag{11}$$

This is what we prove as outlined next.

5) Constrained Perturbation Analysis: The desired inequality (11) is implied by the following condition on U, V: $\|U\|_F \leq \|X\|_F, \|V\|_F \leq \|Y\|_F$ when $\|X\|_F, \|Y\|_F$ are large. Recall that previously we try to find U, V that are close to X, Y; see Proposition 4.1. Now we need to impose extra constraints on U, V, giving rise to Proposition 4.2. The extra constraints make the perturbation analysis significantly more involved; in fact, we apply a sophisticated iterative procedure to construct the factorization $M = UV^T$. The main steps of the proof are briefly given in Appendix C.2.

One crucial component of our proof can be viewed as the perturbation analysis for "preconditioning". Roughly speaking, the basic problem is: given an $r \times r$ matrix $\hat{X}$ with a large condition number, find another matrix $\hat{U}$ with the same Frobenius norm as $\hat{X}$ but smaller inverse Frobenious norm (i.e. $\|\hat{U}^{-1}\|_F \leq \frac{1}{1-\delta}\|\hat{X}^{-1}\|_F$). In other words, we want to reduce $\sum_{i=1}^r \frac{1}{\sigma_i^2}$ with $\sum_{i=1}^r \sigma_i^2$ fixed, where σ_i's are all singular values. Intuitively, by reducing $\sum_{i=1}^r \frac{1}{\sigma_i^2}$ we reduce the discrepancy of singular values. This process is somewhat similar to preconditioning in numerical algebra that reduces the gap between the largest and smallest eigenvalue. The precise statement of the basic problem and its relation with the key technical result Proposition 4.2 are provided in Appendix C.2.2.

6) Algorithm Requirements: We provide three conditions and show that if an algorithm satisfies either of them, then with specific initialization the iterates will stay in the desired basin (see Proposition 5.1). A special case of the third condition has been used in [31] for Grassmann manifold optimization. Together, these three conditions cover a wide spectrum of algorithms including GD, SGD and block coordinate descent type methods.

7) Proof Outline: The overall proof can be divided into two parts: the geometrical property (Lemma 3.1) and the algorithm property (Lemma 3.2). For the geometrical property, Lemma 3.1 states that the regularized objective function $F + G_1 + G_2$ enjoys some nice geometrical property in a certain local region around the global optima, thus there is no other stationary point in this region. For the algorithm property, Lemma 3.2 states that starting from an easily computable initial point, many standard algorithms generate a sequence that are inside the desired region and these algorithms also converge to stationary points. Since these stationary points must be global optima by Lemma 3.1, we obtain that these algorithms converge to the global optima.

D. Other Remarks

1) Difference With Previous Works: As discussed earlier, one major challenge is to bound $\mathcal{P}_\Omega(A)$ when A may be dependent on Ω. One simple strategy as adopted in [17]–[20] is to use a resampling scheme to decouple A and the observation set. This strategy artificially avoids this difficulty, and causes a few issues discussed earlier in Section I-B. Another strategy, as employed in [31], is to use a random graph lemma in [46].

We apply the random graph lemma of [46] when extending the local geometry of $\|M - XY^T\|_F^2$ to $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2$. The difference of our work with [31] is that we study the local geometry in Euclidean space (and, indirectly, the geometry of the quotient manifold), which is quite different from the local geometry in Grassmann manifold studied in [31]. Technically, the complications of the proof in [31] are mostly due to heavy computation of various quantities in Grassmann manifold; in addition, much effort is spent in estimating the terms related to the extra factor S which enables the decoupling of X and Y ([31] actually uses a three-factor decomposition XSY^T). For our problem, one difficulty is to "pull back" the distance in the quotient manifold to the Euclidean space, by the coupled perturbation analysis. Another difficulty is to align the gradient of the regularizer with the global direction (this is not an issue for Grassman manifold), which requires a more sophisticated perturbation analysis. The difficulties have been discussed in detail in Section I-C.

2) Symmetric PSD or Rank-1 Case: The symmetric PSD (positive semi-definite) case or the rank-1 case are easier to deal with, because in the 3-step study of the local geometry the third step is not necessary. When M is rank-1 (possibly non-symmetric), the regularizer $G_2(\cdot)$ may still be needed, but Proposition 4.2 is trivial since its assumptions cannot hold for $r = 1$. When M is symmetric PSD, a popular approach is to use a symmetric factorization $M = XX^T$ instead of the non-symmetric factorization, and the loss function becomes $\|\mathcal{P}_\Omega(M - XX^T)\|_F^2$. The same proof in our paper can be translated to this symmetric PSD case, except that the third step is not necessary. In fact, it is possible to show that (10) holds without any additional requirement on $\boldsymbol{x}$. As a result, the regularizer G_2 and a major technical result Proposition 4.2 are not needed. In both the symmetric PSD and rank-1 case, we only need to establish the intermediate result (7) and the proof can be greatly simplified. Stronger sample complexity and time complexity bounds may be established in these two cases.

3) Simulation Results: The regularizers are introduced due to theoretical purposes; interestingly, they turn out to be helpful in the numerical experiments (the comments below are extracted from the thesis [30, Ch. 2]).

First, the simulation suggests that the imbalance of the rows of X or Y is an important issue for matrix completion in practice, a phenomenon not reported before to our knowledge. The table in [30, Fig. 2.10] shows that when $|\Omega|$ is small, in all successful instances the iterates are balanced, while in all failed instances the iterates are unbalanced. This contrast occurs for many standard algorithms such as AltMin, GD and SGD.

Second, adding only the regularizer G_1 helps, but not too much. Adding an extra regularizer G_2 can push the sample complexity to be very close to the fundamental limit, at least for the synthetic Gaussian data. These experiments seem to indicate that the new regularizers do change the geometry of the problem.

4) Necessity of Incoherence?: While our regularizers are helpful when $|\Omega|$ is small, an open question is whether the row-norm requirement is needed for the local geometry when $|\Omega|$ is large. We observe that the row-norms can be automatically controlled by standard algorithms for the synthetic Gaussian data when there are, say, $5rn$ samples for $n \times n$ matrices. There are two possible explanations (assuming a large $|\Omega|$): (i) the local geometrical property (7) holds without the incoherence requirement; (ii) (7) still requires incoherence, but there is an unknown mechanism for many algorithms to control the row-norms.

To exclude the first possibility, we need to find $(X, Y) \in K(\delta)$ such that $\nabla F(X, Y) = 0$ but $XY^T \neq M$; since (7) holds, such (X, Y) must have unbalanced row-norms. Such an example would validate the necessity of the incoherence restriction for the local geometry. Note that the necessity of incoherence for the local geometry is different from the necessity of an incoherence regularizer/constraint for a specific algorithm. Even if the local geometry requires incoherence, it remains an interesting question why many algorithms can automatically control row-norms when $|\Omega|$ is large.

E. Notations and Organization

1) Notations: Throughout the paper, $M \in \mathbb{R}^{m\times n}$ denotes the unknown data matrix we want to recover, and $r \ll \min\{m, n\}$ is the rank of M. The SVD of M is $M = \hat{U}\Sigma\hat{V}^T$, where $\hat{U} \in \mathbb{R}^{m\times r}, \hat{V} \in \mathbb{R}^{n\times r}$ and $\Sigma \in \mathbb{R}^{r\times r}$ is a diagonal matrix with diagonal entries $\Sigma_1 \geq \Sigma_2 \geq \cdots \geq \Sigma_r$. We denote the maximum and minimum singular value as $\Sigma_{\max}$ and $\Sigma_{\min}$, respectively, and denote $\kappa \triangleq \Sigma_{\max}/\Sigma_{\min}$ as the condition number of M. Define $\alpha = m/n$, which is assumed to be bounded away from 0 and ∞ as $n \longrightarrow \infty$. Without loss of generality, assume $m \geq n$, then $\alpha \geq 1$.

Define the short notations $[m] \triangleq \{1, 2, \ldots, m\}, [n] \triangleq \{1, 2, \ldots, n\}$. Let $\Omega \subseteq [m] \times [n]$ be the set of observed positions, i.e. $\{M_{ij} \mid (i, j) \in \Omega\}$ is the set of all observed entries of M, and define $p \triangleq \frac{|\Omega|}{mn}$ which can be viewed as the probability that each entry is observed. For a linear subspace $\mathcal{S}$, denote $\mathcal{P}_\mathcal{S}$ as the projection onto $\mathcal{S}$. By a slight abuse of notation, we denote $\mathcal{P}_\Omega$ as the projection onto the subspace $\{W \in \mathbb{R}^{m\times n} : W_{i,j} = 0, \forall (i, j) \notin \Omega\}$. In other words, $\mathcal{P}_\Omega(A)$ is a matrix where the entries in Ω are the same as A while the entries outside of Ω are zero.

For a vector $x \in \mathbb{R}^n$, denote $\|x\|$ as its Euclidean norm. For a matrix X, denote $\|X\|_F$ as its Frobenius norm, and $\|X\|_2$ as its spectral norm (i.e. the largest singular value). Denote $\sigma_{\max}(X), \sigma_{\min}(X)$ as the largest and smallest singular values of X, respectively. Let $X^\dagger$ denote the pseudo inverse of a matrix X. The standard inner product between vectors or matrices are written as $\langle x, y\rangle$ or $\langle X, Y\rangle$, respectively. Denote $A^{(i)}$ as the ith row of a matrix A. We will use C, C_1, C_T, C_d, etc. to denote universal numerical constants.

2) Organization: The rest of the paper is organized as follows. In Section II we introduce the problem formulation and four typical algorithms. In Section III, we present the main results and the main lemmas used in the proofs of these results. The proof of the two lemmas used in proving Theorem 3.1 are given in Section IV and Section V respectively. The proof of the first lemma depends on two "coupled perturbation analysis" results Proposition 4.1 and Proposition 4.2, the proofs of which are given in Appendix A.2 and Appendix B.1 respectively. The proof of a lemma used in proving Theorem 3.2 is given in Appendix D.5.

II. Problem Formulation and Algorithms

A. Assumptions

1) Incoherence Condition: The incoherence condition for the matrix completion problem is first introduced by Candès and Recht in [4] and has become a standard assumption for low-rank matrix recovery problems (except a few recent works such as [47] and [48]). We will define an incoherence condition for an $m \times n$ matrix M which is the same as that in [31].

Definition 2.1: We say a matrix $M = \hat{U}\Sigma\hat{V}^T$ (compact SVD of M) is μ-incoherent if:

$$\sum_{k=1}^{r} \hat{U}_{ik}^2 \le \frac{\mu r}{m}, \quad \sum_{k=1}^{r} \hat{V}_{jk}^2 \le \frac{\mu r}{n}, \ 1 \le i \le m, 1 \le j \le n. \tag{12}$$

It can be shown that $\mu \in [1, \frac{\max\{m,n\}}{r}]$. For some popular random models for generating M, the incoherence condition holds with a parameter scaling as $\sqrt{r \log n}$ (see [31]). In this paper, we just assume that M is μ-incoherent. Note that the incoherence condition implies that $\hat{U}, \hat{V}$ have bounded row norm. Throughout the paper, we also use the terminology "incoherent" to (imprecisely) describe $m \times r$ or $n \times r$ matrices that have bounded row norm (see the definition of set K_1 in (30)).

2) Random Sampling Model: In the statement of the results in this paper, the probability is taken with respect to the uniform random model of $\Omega \subseteq [m] \times [n]$ with fixed size $|\Omega| = S$ (i.e. Ω is generated uniformly at random from set $\{\Omega' \subseteq [m] \times [n] : \text{ the size of } \Omega' \text{ is } S\}$). We remark that this model is "equivalent to" a Bernolli model that each entry of M is included into Ω independently with probability $p = \frac{S}{mn}$ in the sense that if the success of an algorithm holds for the Bernolli model with a certain p with high probability, then the success also holds for the uniform random model with $|\Omega| = pmn$ with high probability (see [4] or [31, Sec. 1D] for more details). Thus in the proofs we will instead use the Bernolli model.

B. Problem Formulation

We consider a variant of (P0) with incoherence-control regularizers. In particular, we introduce two types of regularization terms besides the square loss function: the first type is designed to force the iterates X_k, Y_k to be incoherent (i.e. with bounded row norm), and the second type is designed to upper bound the norm of X_k and Y_k. Note that (P0) is related to the Lagrangian method, while our regularizer is based on the penalty function method for constrained optimization problems. We can also view the regularizer $\lambda(\|X\|_F^2 + \|Y\|_F^2)$ as a "soft regularizer", and our new regularizer as a "hard regularizer". The advantage of the hard regularizer is that it does not distort the optimal solution.

Our regularizers are smooth functions with simple gradients, thus the algorithms for our formulation have similar per-iteration computation cost as the algorithms for the formulation without regularizers. In the numerical experiments, we find that when $|\Omega|$ is large, the iterates are always incoherent and bounded, and our algorithms are the same as the traditional algorithms for the unregularized formulation; when $|\Omega|$ is relatively small, the traditional algorithms may produce high error, and our regularizer becomes active and significantly reduce the error. In some sense, our algorithms for the new formulation are "better" versions of the traditional algorithms, and our theoretical results can be viewed as a validation of the traditional algorithms in the "large-$|\Omega|$ regime" and a validation of the modified algorithm in the "small-Ω" regime. Preliminary simulation results show that many algorithms for the proposed formulation can recover the matrix when $|\Omega|$ is very close to the fundamental limit, significantly improving upon the traditional algorithms; see [30, Ch. 3].

The regularization function G is defined as follows:

$$\begin{aligned} G(X,Y) \triangleq \rho \sum_{i=1}^{m} G_0\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right) + \rho \sum_{j=1}^{n} G_0\left(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2}\right) \\ + \rho G_0\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right) + \rho G_0\left(\frac{3\|Y\|_F^2}{2\beta_T^2}\right), \end{aligned} \tag{13}$$

where $A^{(i)}$ denotes the ith row of a matrix A,

$$G_0(z) \triangleq I_{[1,\infty]}(z)(z-1)^2 = \max\{0, z-1\}^2, \tag{14}$$

$$\begin{aligned} \beta_T &\triangleq \sqrt{C_T r \Sigma_{\max}}, \ \beta_1 \triangleq \beta_T\sqrt{\frac{3\mu r}{m}} = \sqrt{C_T r \Sigma_{\max}}\sqrt{\frac{3\mu r}{m}}, \\ \beta_2 &\triangleq \beta_T\sqrt{\frac{3\mu r}{n}} = \sqrt{C_T r \Sigma_{\max}}\sqrt{\frac{3\mu r}{n}}. \end{aligned} \tag{15}$$

Here, $I_{\mathcal{C}}$ is the indicator function of a set $\mathcal{C}$, i.e. $I_{\mathcal{C}}(z)$ equals 1 when $z \in \mathcal{C}$ and 0 otherwise. ρ is a constant specified shortly. Throughout the paper, δ and δ_0 are defined as

$$\delta \triangleq \frac{\Sigma_{\min}}{C_d r^{1.5}\kappa}, \quad \delta_0 \triangleq \frac{\delta}{6}, \tag{16}$$

where C_d is some numerical constant. The coefficient ρ is defined as (a larger ρ also works)

$$\rho \triangleq \frac{2p\delta_0^2}{G_0(3/2)} = 8p\delta_0^2. \tag{17}$$

The numerical constant $C_T > 5$ will be specified in the proof of our main result. The parameter β_T is chosen to be of the same order as $\|\hat{U}\Sigma^{1/2}\|_F$ and $\|\hat{V}\Sigma^{1/2}\|_F$, and β_1, β_2 are chosen to be of the same order as $\sqrt{r}\|(\hat{U}\Sigma^{1/2})^{(i)}\|, \sqrt{r}\|(\hat{V}\Sigma^{1/2})^{(j)}\|$. The additional factor $\sqrt{3r}$ is due to technical consideration (to prove (256)). Our regularizer G involves $\Sigma_{\max}$ and μ which depend on the unknown matrix M; in practice, we can estimate $\Sigma_{\max}$ by $c_1\sqrt{\frac{\|\mathcal{P}_\Omega(M)\|_F^2}{pr}}$, and estimate μ by $c_2\frac{\sqrt{mn}}{r\Sigma_{\max}}\max_{(i,j)\in\Omega}|M_{ij}|$ (according to (203)) where c_1, c_2 are numerical constants to tune.

It is easy to verify that G_0 is continuously differentiable. The choice of function G_0 is not unique; in fact, we can choose any G_0 that satisfies the following requirements: a) G_0 is convex and continuously differentiable; b) $G_0(z) = 0, z \in [0, 1]$. In [31], G_0 is chosen as $G_0(z) = I_{[1,\infty]}(z)(e^{(z-1)^2} - 1)$, which also satisfies these two requirements. Choosing different G_0 does not affect the proof except the change of numerical constants (which depend on $G_0(3/2), G_0'(3/2), G_0''(3/2)$). Note that the requirement of G_0 being non-decreasing and convex guarantees the convexity of $G(X,Y)$. In fact, according to the well-known result that the composition of a non-decreasing convex function and a convex function is a convex function, and notice that $\|X^{(i)}\|^2, \|Y^{(j)}\|^2, \|X\|_F^2, \|Y\|_F^2$ are convex, we have that each component of G is convex and thus G is convex.

Denote the square loss term in (P0) as $F(X,Y) \triangleq \sum_{(i,j)\in\Omega}[M_{ij} - (XY^T)_{ij}]^2 = \|\mathcal{P}_\Omega(M - XY^T)\|_F^2$. Replacing the objective function of (P0) by $\tilde{F}(X,Y) \triangleq F(X,Y) + G(X,Y)$, we obtain the following problem:

$$\text{P1}: \min_{X\in\mathbb{R}^{m\times r}, Y\in\mathbb{R}^{n\times r}} \frac{1}{2}\|\mathcal{P}_\Omega(M - XY^T)\|_F^2 + G(X,Y). \tag{18}$$

We remark that (P1) can be interpreted as the penalized version of the following constrained problem (see, e.g. [49])

$$\begin{aligned}
&\min_{X,Y} \ \frac{1}{2}\|\mathcal{P}_\Omega(M - XY^T)\|_F^2,\\
&\text{s.t.}\ \|X\|_F^2 \le \frac{2}{3}\beta_T^2, \quad \|Y\|_F^2 \le \frac{2}{3}\beta_T^2;\\
&\qquad \|X^{(i)}\|^2 \le \frac{2}{3}\beta_1^2, \quad \forall i, \ \|Y^{(j)}\|^2 \le \frac{2}{3}\beta_2^2, \ \forall j.
\end{aligned} \tag{19}$$

To illustrate this, note that the constraint $f_1(X) \triangleq \frac{3\|X\|_F^2}{2\beta_T^2} - 1 \le 0$ corresponds to the penalty term $\rho G_0(f_1(X) + 1) = \rho\max\{0, f_1(X)\}^2$ which appears as the third term in $G(X,Y)$, and similarly other constraints correspond to other terms in $G(X,Y)$. In other words, the regularization function $G(X,Y)$ is just a penalty function for the constraints of the problem (19). The function $\max\{0,\cdot\}^2$ is a popular choice for the penalty function in optimization (see, e.g. [49]), which motivates our choice of G_0 in (14). Our result can be extended to cover the algorithms for the constrained version (19), or a partially regularized formulation (e.g. only penalize the violation of the constraint $\|X\|_F^2 \le \frac{2}{3}\beta_T^2, \|Y\|_F^2 \le \frac{2}{3}\beta_T^2$).

It is easy to check that the optimal value of (P1) is zero and $(X,Y) = (\hat{U}\Sigma^{1/2}, \hat{V}\Sigma^{1/2})$ is an optimal solution to (P1), provided that M is μ-incoherent. In fact, since $\tilde{F}$ is a nonnegative function, we only need to show $\tilde{F}(X,Y) = 0$ for this choice of (X,Y). As $XY^T = M$ implies $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2 = 0$, we only need to show $G(X,Y) = G(\hat{U}\Sigma^{1/2}, \hat{V}\Sigma^{1/2})$ equals zero. In the expression of $G(X,Y)$, the third and fourth terms $G_0(\frac{3\|X\|_F^2}{2\beta_T^2})$ and $G_0(\frac{3\|Y\|_F^2}{2\beta_T^2})$ equal zero because $\|X\|_F^2 = \|Y\|_F^2 \le r\Sigma_{\max} < \frac{2}{3}\beta_T^2$. The first and second terms $\sum_i G_0(\frac{3\|X^{(i)}\|^2}{2\beta_1^2})$ and $\sum_j G_0(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2})$ equal zero because $\|X^{(i)}\|^2 \le \Sigma_{\max}\|\hat{U}^{(i)}\|^2 \le \Sigma_{\max}\frac{\mu r}{m} \le \frac{2}{3}\beta_1^2$, for all i and, similarly, $\|Y^{(j)}\|^2 \le \frac{2}{3}\beta_2^2$, for all j, where we have used the incoherence condition (12). This verifies our previous claim that the "hard regularizer" $G(X,Y)$ does not distort the optimal solution of the original formulation.

One commonly used assumption in the optimization literature is that the gradient of the objective function is Lipschitz continuous. For any positive number β, define a bounded set

$$\begin{aligned}\Gamma(\beta) \triangleq \{(X,Y) | X \in \mathbb{R}^{m\times r}, Y \in \mathbb{R}^{n\times r},\\
\|X\|_F \le \beta, \|Y\|_F \le \beta\}.\end{aligned} \tag{20}$$

The following result shows that this assumption (Lipschitz continuous gradients) holds for our objective function within a bounded set.

Claim 2.1: Suppose $\beta_0 \ge \beta_T$ and

$$L(\beta_0) \triangleq 4\beta_0^2 + 54\rho\frac{\beta_0^2}{\beta_1^4}. \tag{21}$$

Then $\nabla\tilde{F}(X,Y)$ is Lipschitz continuous over the set $\Gamma(\beta_0)$ with Lipschitz constant $L(\beta_0)$, i.e.

$$\begin{aligned}&\|\nabla\tilde{F}(X,Y) - \nabla\tilde{F}(U,V)\|_F \le L(\beta_0)\|(X,Y) - (U,V)\|_F,\\
&\qquad \forall (X,Y), (U,V) \in \Gamma(\beta_0),\end{aligned}$$

where $\|(X,Y) - (U,V)\|_F = \sqrt{\|X-U\|_F^2 + \|Y-V\|_F^2}$.

The proof of Claim 2.1 is given in Appendix A.1.

C. Row-Scaled Spectral Initialization

Our results require the initial point to be close enough to the global optima. To be more precise, we want the initial point to be in an incoherent neighborhood of the original matrix M (this neighborhood will be specified later). Special initialization is also required in other works on non-convex formulations [17]–[20], [31], [32], [36].

We will show that such an initial point can be found through a simple procedure. This procedure consists of two steps: first, using the spectral method (see, e.g. [31]), we obtain $M_0 = \hat{X}_0\hat{Y}_0^T$ which is close to M; second, we scale the rows of $(\hat{X}_0, \hat{Y}_0)$ to make it incoherent (i.e. with bounded row-norm). Denote the best rank-r approximation of a matrix A as $\mathrm{P}_r(A)$. Define an operation SVD_r that maps a matrix A to the SVD

TABLE I

INITIALIZATION PROCEDURE (INITIALIZE)

Input: $\mathcal{P}_\Omega(M)$, target rank r, target row norm bounds β_1, β_2.

Algorithm INITIALIZE($\mathcal{P}_\Omega(M), p, r$).

1. Compute $(\bar{X}_0, D_0, \bar{Y}_0) = \mathrm{SVD}_r\left(\frac{1}{p}\mathcal{P}_\Omega(M)\right)$, as defined in (22).
   Compute $\hat{X}_0 = \bar{X}_0 D_0^{1/2}, \hat{Y}_0 = \bar{Y}_0 D_0^{1/2}$.

2. For each row of $\hat{X}_0$ (resp. $\hat{Y}_0$) with norm larger than $\sqrt{\frac{2}{3}}\beta_1$ (resp. $\sqrt{\frac{2}{3}}\beta_2$), scale it to make the norm of this row equal $\sqrt{\frac{2}{3}}\beta_1$ (resp. $\sqrt{\frac{2}{3}}\beta_2$) to obtain X_0, Y_0, i.e.

$$\begin{aligned} X_0^{(i)} &= \frac{\hat{X}_0^{(i)}}{\|\hat{X}_0^{(i)}\|} \min\left\{\|\hat{X}_0^{(i)}\|, \sqrt{\frac{2}{3}}\beta_1\right\}, i = 1, \ldots, m. \\ Y_0^{(j)} &= \frac{\hat{Y}_0^{(j)}}{\|\hat{Y}_0^{(j)}\|} \min\left\{\|\hat{Y}_0^{(j)}\|, \sqrt{\frac{2}{3}}\beta_2\right\}, j = 1, \ldots, n. \end{aligned} \tag{23}$$

Output $X_0 \in \mathbb{R}^{m\times r}, Y_0 \in \mathbb{R}^{n\times r}$.

TABLE II

ALGORITHM 1 (GRADIENT DESCENT)

Initialization: $(X_0, Y_0) \longleftarrow$ INITIALIZE($\mathcal{P}_\Omega(M), p, r$).

The k-th iteration:

$$X_k \longleftarrow X_k(\eta_k) \triangleq X_{k-1} - \eta_k \nabla_X \tilde{F}(X_{k-1}, Y_{k-1}),$$
$$Y_k \longleftarrow Y_k(\eta_k) \triangleq Y_{k-1} - \eta_k \nabla_Y \tilde{F}(X_{k-1}, Y_{k-1}),$$

where the stepsize η_k is chosen according to one of the following rules:

a) Constant stepsize: $\eta_k = \eta \le \bar{\eta}_1, \ \forall\ k$ ($\bar{\eta}_1$ is a constant defined by (238) in Appendix L).

b) Restricted Armijo rule: Let $\sigma \in (0,1), \xi \in (0,1), s_0$ be fixed scalars.

b1) Find the smallest nonnegative integer i such that $d(\boldsymbol{x}_k(\xi^i s_0), \boldsymbol{x}_0) \le 5\delta/6$ and $\tilde{F}(\boldsymbol{x}_k(\xi^i s_0)) \le \tilde{F}(\boldsymbol{x}_{k-1}) - \sigma\xi^i s_0 \|\nabla \tilde{F}(\boldsymbol{x}_{k-1})\|_F^2$.

b2) Let $\eta_k = \xi^i s_0$.

c) Restricted line search: $\eta_k = \arg\min_{\eta\in\mathbb{R}, d(\boldsymbol{x}_k(\eta),\boldsymbol{x}_0)\le 5\delta/6} \tilde{F}(\boldsymbol{x}_k(\eta))$.

components (X, D, Y) of its best rank-r approximation $\mathrm{P}_r(A)$, i.e.

$$\mathrm{SVD}_r(A) \triangleq (X, D, Y), \quad \text{where } XDY^T \text{ is compact SVD of } \mathrm{P}_r(A). \tag{22}$$

The initialization procedure is given in Table I. The property of the initial point generated by this procedure will be presented in Claim 5.2.

In the numerical experiments, we find that the proposed initialization is not better than random initialization if we use the proposed formulation with the incoherence-control regularizer. In contrast, for traditional formulations (either unregularized or with a regularizer $\lambda(\|X\|_F^2 + \|Y\|_F^2)$) the proposed initialization does lead to better recovery performance (lower sample complexity). We also notice that the row-scaling step is crucial for this improvement since simply initializing via the spectral method does not help too much. See [30, Ch. 3] for the simulation results and discussions.

D. Algorithms

Our result applies to many standard algorithms such as gradient descent, SGD and block coordinate descent type methods (including alternating minimization, block coordinate gradient descent, block successive upper bound minimization, etc.). We will describe several typical algorithms in this subsection.

The gradient $\nabla\tilde{F} = \nabla F + \nabla G = (\nabla_X F + \nabla_X G, \nabla_Y F + \nabla_Y G)$ can be easily computed as follows:

$$\begin{aligned} \nabla_X F(X,Y) &= \mathcal{P}_\Omega(XY^T - M)Y, \\ \nabla_Y F(X,Y) &= \mathcal{P}_\Omega(XY^T - M)^T X, \\ \nabla_X G(X,Y) &= \rho \sum_{i=1}^m G_0'\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right)\frac{3\bar{X}^{(i)}}{\beta_1^2} \\ &\quad + \rho G_0'\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right)\frac{3X}{\beta_T^2}, \\ \nabla_Y G(X,Y) &= \rho \sum_{j=1}^n G_0'\left(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2}\right)\frac{3\bar{Y}^{(j)}}{\beta_2^2} \\ &\quad + \rho G_0'\left(\frac{3\|Y\|_F^2}{2\beta_T^2}\right)\frac{3Y}{\beta_T^2}, \end{aligned} \tag{24}$$

where $G_0'(z) = I_{[1,\infty]}(z)2(z-1)$, and $\bar{X}^{(i)}$ (resp. $\bar{Y}^{(j)}$) denotes a matrix with the i-th (resp. j-th) row being $X^{(i)}$ (resp. $Y^{(j)}$) and the other rows being zero.

We first present a gradient descent algorithm in Table II. There are many choices of stepsizes such as constant stepsize, exact line search, limited line search, diminishing stepsize and Armijo rule [50]. We present three stepsize rules here: constant stepsize, restricted Armijo rule and restricted line search (the latter two are the variants of Armijo rule and exact line search). Note that the restricted line search rule is similar to that used in [31] for the gradient descent method

TABLE III

ALGORITHM 2 (TWO-BLOCK ALTERNATING MINIMIZATION)

Initialization: $(X_0, Y_0) \longleftarrow$ INITIALIZE$(\mathcal{P}_\Omega(M), p, r)$.
The k-th iteration:
$X_k \longleftarrow \arg\min_X \tilde{F}(X, Y_{k-1})$,
$Y_k \longleftarrow \arg\min_Y \tilde{F}(X_{k-1}, Y)$.

over Grassmannian manifolds. To simplify the notations, we denote $\boldsymbol{x}_k(\eta) \triangleq (X_k(\eta), Y_k(\eta))$ and $d(\boldsymbol{x}_k(\eta), \boldsymbol{x}_0) \triangleq \sqrt{\|X_k(\eta) - X_0\|_F^2 + \|Y_k(\eta) - Y_0\|_F^2}$.

AltMin (alternating minimization) belongs to the class of block coordinate descent (BCD) type methods. One can update the blocks in different orders (e.g. cyclic [51]–[53], randomized [54] or parallel) and solve the subproblem inexactly. Commonly used inexact BCD type algorithms include BCGD (block coordinate gradient descent, which updates each variable by a single gradient step [54]) and BSUM (block successive upper bound minimization, which updates each variable by minimizing an upper bound of the objective function [55]). BCD-type methods have been widely used in engineering (e.g. [56], [57]). In the context of matrix completion, Hastie et al. [58] proposed an algorithm that could be viewed as a BSUM algorithm. Just considering different choices of the blocks will lead to different algorithms for the matrix completion problem [29]. Our result applies to many BCD type methods, including the two-block alternating minimization, BCGD and BSUM. While it is not very interesting to list all possible algorithms to which our results are applicable, we just present two specific algorithms for illustration.

The first BCD type algorithm we present is (two-block) AltMin, which, in the context of matrix completion, usually refers to the algorithm that alternates between X and Y by updating one factor at a time with the other factor fixed. Although the overall objective function is non-convex, each subproblem of X or Y is convex and thus can be solved efficiently. The details are given in Table III.

For the case without the regularization term $G(X, Y)$, the objective function becomes $F(X, Y)$ and is quadratic with respect to X or Y. Thus X_k, Y_k have closed form update. Suppose $X^T = (x_1, \ldots, x_m)$ and $Y^T = (y_1, \ldots, y_n)$, where $x_i, y_j \in \mathbb{R}^{r\times 1}$. Then $(x_1^*, \ldots, x_m^*) \triangleq (\arg\min_X F(X, Y))^T$ and $(y_1^*, \ldots, y_n^*) \triangleq (\arg\min_Y F(X, Y))^T$ are given by

$$\begin{aligned} x_i^* &= (\sum_{j\in\Omega_i^x} y_j y_j^T)^\dagger (\sum_{j\in\Omega_i^x} M_{ij} y_j), \quad i = 1, \ldots, m, \\ y_j^* &= (\sum_{i\in\Omega_j^y} x_i x_i^T)^\dagger (\sum_{i\in\Omega_j^y} M_{ij} x_i), \quad j = 1, \ldots, n, \end{aligned} \tag{25}$$

where $\Omega_i^x = \{j \mid (i, j) \in \Omega\}, \Omega_j^y = \{i \mid (i, j) \in \Omega\}$, and $A^\dagger$ denotes the pseudo inverse of a matrix A. For our problem with the regularization term $G(X, Y)$, we no longer have closed form update of X_k, Y_k. One way to solve the convex subproblems is to start from the solution given in (25) and then apply the gradient descent method until convergence. The details for solving $\min_X \tilde{F}(X, Y)$ is given in Table IV (the stepsize can be chosen by one of the standard rules of the gradient descent method), and the other subproblem $\min_Y \tilde{F}(X, Y)$ can be solved in a similar fashion.

Theoretically speaking, AltMin for our formulation (P1) is not as efficient as the vanilla AtlMin for (P0) since an extra inner loop is needed to solve the subproblem. However, we remark that in the regimes of $|\Omega|$ that the vanilla AltMin works, the least square solution X (resp. Y) is always bounded and incoherent (empirical observation), in which case the regularizer G is inactive; therefore, the gradient updates in Table IV do not happen. In the regimes of $|\Omega|$ that the vanilla AltMin fails, G is active and the gradient updates do happen; however, instead of solving the subproblem exactly, one could perform one gradient step and the algorithm becomes the popular variant BCGD [54]. Our main result of exact recovery still holds for BCGD (the proof for Algorithm 3 in Claim 5.3 can be applied to BCGD since BCGD is a special case of BSUM).

In the second BCD type algorithm called row BSUM, we update the rows of X and Y cyclically by minimizing an upper bound of the objective function; see Table V. The extra terms $\frac{\lambda_0}{2}\|X^{(i)} - X_{k-1}^{(i)}\|^2$ or $\frac{\lambda_0}{2}\|Y^{(j)} - Y_{k-1}^{(j)}\|^2$ are added to make the subproblems strongly convex, which help prove convergence to stationary points. Such a technique has also been used in the alternating least square algorithm for tensor decomposition [55]. Note that for the two-block BCD algorithm, convergence to stationary points can be guaranteed even when the subproblems are not strongly convex [59], thus in Algorithm 2 we do not add the extra terms. The benefit of cyclically updating the rows is that each subproblem can be solved efficiently using a simple binary search; see Appendix A.2 for the details. We remark again that instead of solving the subproblem exactly, one could just perform one gradient step to update each row of X and Y (with $\lambda = 0$) and our result still holds.

The fourth algorithm we present is SGD (stochastic gradient descent) [1], [23] tailored for our problem (P1). In the optimization literature, this algorithm for minimizing the sum of finitely many functions is more commonly referred to as “incremental gradient method”, while SGD represents the algorithm for minimizing the expectation of a function; nevertheless, in this paper we follow the convention in the computer science literature and still call it “SGD”. In SGD, at each iteration we pick a component function and perform a gradient update. Similar to the BCD type methods where the blocks can be chosen in different orders, one can pick the component functions in a cyclic order, in an essentially cyclic order, or in a random order (either sampling with replacement or without replacement). In practice, the version of sampling without replacement converges much faster than the version of sampling with replacement (see [30, Ch. 2] for simulation results). In general, the understanding of sampling without replacement for optimization algorithms is quite limited (see, e.g., [60] for one example of such analysis).

In this paper we only consider the cyclic order, and use a standard stepsize rule for SGD [61], [62] which requires the stepsizes $\{\eta_k\}$ to go to zero as $k \to \infty$, but neither too fast nor too slow (this choice guarantees convergence to stationary

TABLE IV
SOLVING SUBPROBLEM OF ALGORITHM 2

Solving subproblem of Algorithm 2: $\min_X \tilde{F}(X, Y)$.
Input: $Y = (y_1, \dots, y_n) \in \mathbb{R}^{n\times r}$.
Initialization: $X = (x_1, \dots, x_m)$, where $x_i = (\sum_{j\in\Omega_i^x} y_j y_j^T)^\dagger(\sum_{j\in\Omega_i^x} M_{ij} y_j)$, $i = 1, \dots, m$,
Repeat:
$X \longleftarrow X - \eta\nabla_X \tilde{F}(X, Y)$,
Until Stopping criterion is met.

TABLE V
ALGORITHM 3 (ROW BSUM)

Initialization: $(X_0, Y_0) \longleftarrow$ INITIALIZE$(\mathcal{P}_\Omega(M), p, r)$.
Parameter: $\lambda_0 > 0$.
The k-th loop:
For $i = 1$ to m:
$X_k^{(i)} \longleftarrow \arg\min_{X^{(i)}} \tilde{F}(X_k^{(1)}, \dots, X_k^{(i-1)}, X^{(i)}, X_{k-1}^{(i+1)} \dots, X_{k-1}^{(m)}, Y_{k-1}) + \frac{\lambda_0}{2}\|X^{(i)} - X_{k-1}^{(i)}\|^2$,
For $j = 1$ to n:
$Y_k^{(j)} \longleftarrow \arg\min_{Y^{(j)}} \tilde{F}(X_k, Y_k^{(1)}, \dots, Y_k^{(j-1)}, Y^{(j)}, Y_{k-1}^{(j+1)}, \dots, Y_{k-1}^{(m)}) + \frac{\lambda_0}{2}\|Y^{(j)} - Y_{k-1}^{(j)}\|^2$.

TABLE VI
ALGORITHM 4 (SGD)

Initialization: $(X_0, Y_0) \longleftarrow$ INITIALIZE$(\mathcal{P}_\Omega(M), p, r)$.
Parameters: $\eta_k, k = 0, 1, \dots$ satisfying $\sum_k \eta_k = \infty, \sum_k \eta_k^2 < \eta_{\text{sum}}$ and $0 < \eta_k \le \bar{\eta}$, where η_{sum} and $\bar{\eta}$ are constants specified in Appendix L.
The $(k+1)$-th loop:
$X_{k,0} \longleftarrow X_k, \quad Y_{k,0} \longleftarrow Y_k$.
For $i = 1$ to $\lvert\Omega\rvert + m + n + 2$:
$X_{k,i} \longleftarrow X_{k,i-1} - \eta_k \nabla_X f_i(X_{k,i-1}, Y_{k,i-1})$,
$Y_{k,i} \longleftarrow Y_{k,i-1} - \eta_k \nabla_Y f_i(X_{k,i-1}, Y_{k,i-1})$.
End
$X_{k+1} \longleftarrow X_{k,\lvert\Omega\rvert+m+n+2}, \quad Y_{k+1} \longleftarrow Y_{k,\lvert\Omega\rvert+m+n+2}$.

points even for nonconvex problems). One such choice of stepsizes is $\eta_k = O(1/k)$. We remark that our results also apply to other versions of SGD with different update orders or stepsize rules as long as they converge to stationary points.

To apply SGD to our problem, we decompose the objective function $\tilde{F}(X, Y)$ as follows:

$$\begin{aligned}\tilde{F}(X, Y) &= \sum_{(i,j)\in\Omega} F_{ij}(X, Y) + \sum_{i=1}^{m} G_{1i}(X) \\ &\quad + \sum_{j=1}^{n} G_{2j}(Y) + G_3(X) + G_4(Y) \\ &= \sum_{k=1}^{|\Omega|+m+n+2} f_k(X, Y),\end{aligned}$$

where the component functions

$$\begin{aligned}F_{ij}(X, Y) &= [(XY^T - M)_{ij}]^2 \\ &= [(X^{(i)})^T Y^{(j)} - M_{ij}]^2, \quad (i, j) \in \Omega, \\ G_{1i}(X) &= \rho G_0(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}), \quad 1 \le i \le m, \\ G_{2j}(Y) &= \rho G_0(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2}), \quad 1 \le j \le n, \\ G_3(X) &= \rho G_0(\frac{3\|X\|_F^2}{2\beta_T^2}), \\ G_4(Y) &= \rho G_0(\frac{3\|Y\|_F^2}{2\beta_T^2}) \end{aligned} \tag{26}$$

and $\{f_k(X, Y)\}_{k=1}^{|\Omega|+m+n+2}$ denotes the collection of all component functions. With these definitions, the SGD algorithm is given in Table VI.

III. MAIN RESULTS

The main result of this paper is that Algorithms 1-4 (standard optimization algorithms) will converge to the global optima of problem (P1) given in (18) and reconstruct M exactly with high probability, provided that the number of revealed entries is large enough. Similar to the results for nuclear norm minimization [4]–[7], the probability is taken with respect to the random choice of Ω, and the result also applies to a uniform random model of Ω.

Theorem 3.1 (Exact Recovery): Assume a rank-r matrix $M \in \mathbb{R}^{m\times n}$ is μ-incoherent. Suppose the condition number of M is κ and $\alpha = m/n \ge 1$. Then there exists a numerical constant C_0 such that: if Ω is uniformly generated at random with size

$$|\Omega| \ge C_0 \alpha n r \kappa^2 \max\{\mu \log n, \sqrt{\alpha}\mu^2 r^6 \kappa^4\}, \tag{27}$$

then with probability at least $1-2/n^4$, each of Algorithms 1-4 reconstructs M exactly. Here, we say an algorithm reconstructs M if each limit point (X^*, Y^*) of the sequence $\{X_k, Y_k\}$ generated by this algorithm satisfies $X^*(Y^*)^T = M$.

This result shows that although (18) is a non-convex optimization problem, many standard algorithms can converge to the global optima with certain initialization. Different from all previous works on alternating minimization for matrix completion, our result does not require the algorithm to use independent samples in different iterations. To the best of our knowledge, our result is the first one that provides theoretical guarantee for alternating minimization without resampling. In addition, this result also provides the first *exact* recovery guarantee for many algorithms such as gradient descent, SGD and BSUM.

As demonstrated in [4] (and proved in [5, Th. 1.7]), $O(nr \log n)$ entries are the minimum requirement to recover the original matrix: $O(nr)$ is the number of degrees of freedom of a rank r matrix M, and the additional $\log n$ factor is due to the coupon collector effect [4]. For $r = O(1)$ and κ bounded, Theorem 3.1 is order optimal in terms of the sample complexity since only $O(n \log n)$ entries are needed to exactly recover M. For $r = \mathcal{O}(\log n)$, however, our result is suboptimal by a polylogarithmic factor. The initialization has contributed $r^4\kappa^4$ to the sample complexity bound, and we expect that using other initialization procedures (e.g. the one proposed in [19]) can reduce the exponents of r and κ.

Theorem 3.1 only establishes the convergence, but not the convergence speed. With some extra effort, we can prove the linear convergence of the gradient descent method (see Theorem 3.2 below). Again, this result can be extended beyond the gradient descent method. In fact, by a standard optimization argument, we can prove the linear convergence of any algorithm that satisfies "sufficient decrease" (i.e. $\tilde{F}(\boldsymbol{x}^k) - \tilde{F}(\boldsymbol{x}^{k+1}) \geq O(\|\nabla \tilde{F}(\boldsymbol{x}^k)\|_F^2)$) and the requirements in Lemma 3.2; see Corollary 3.2. Many first order methods, including alternating type methods (e.g. BCGD, two-block BCD), can be shown to have the sufficient decrease property under mild conditions. For space reason, we do not verify all the methods considered in this paper, but only present the linear convergence result for the gradient descent method. The proof of Theorem 3.2 is given in Section III-B.

Theorem 3.2 (Linear Convergence): Under the same condition of Theorem 3.1, with probability at least $1 - 2/n^4$, *Algorithm 1a (gradient descent with constant stepsize) converges linearly; more precisely, the sequence* $\{X_k, Y_k\}$ *generated by Algorithm 1a satisfies*

$$\tilde{F}(X_k, Y_k) \leq (1 - \frac{1}{2}\eta_1 \xi)^k, \tag{28}$$

where $\xi = \frac{1}{C_g r^5 \kappa^3} p \Sigma_{\min}$ *(here* C_g *is a numerical constant),* η_1 *is the stepsize and* $\eta_1 \xi < 1$.

The linear convergence will immediately lead to a time complexity of $\tilde{O}(\text{poly}(n) \log \frac{1}{\epsilon})$ for achieving any ϵ-optimal solution, where the $\tilde{O}$ notation hides factors polynomial in r, κ, α. We conjecture that the time complexity bound can be improved to $\tilde{O}(|\Omega| \log(1/\epsilon))$ as observed in practice. However, finding the optimal time complexity bound is not the focus of this paper, and is left as future work.

The above result shows that $\tilde{F}(X_k, Y_k)$ converges to zero at a linear speed. Note that $\tilde{F}(X, Y) = 0$ (global convergence) only implies $\mathcal{P}_\Omega(M - XY^T) = 0$, not necessarily $M = XY^T$ (exact reconvery). The following lemma implies that with high probability (for random Ω) the global convergence implies the exact recovery. In fact, it shows that the observed loss $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2$ is on the order of the recovery error $p\|M - XY^T\|_F^2$ if (X, Y) lies in an incoherent neighborhood of M. As discussed in the introduction, this lemma can also be viewed as a geometrical property of $f_\Omega(Z) = \|\mathcal{P}_\Omega(M - Z)\|_F^2$ in a local incoherent region (view $\mathcal{P}_\Omega(Z - M)$ as the gradient of $f_\Omega(Z)$).

Claim 3.1: Under the same condition of Theorem 3.1, with probability at least $1 - 1/(2n^4)$, *we have*

$$\frac{1}{3} p\|M - XY^T\|_F^2 \leq \|\mathcal{P}_\Omega(M - XY^T)\|_F^2 \leq 2p\|M - XY^T\|_F^2, \quad \forall (X, Y) \in K_1 \cap K_2 \cap K(\delta). \tag{29}$$

The proof of this claim is given in Appendix D.2. This result is a simple corollary of several intermediate bounds established in the proof of Lemma 3.1.

A. Proof of Theorem 3.1 and Main Lemmas

To prove Theorem 3.1, we only need to prove two lemmas which describe the local geometry of the regularized objective in (P1) and the properties of the algorithms respectively. Roughly speaking, the first lemma shows that any stationary point of (P1) in a certain region is globally optimal, and the second lemma shows that each of Algorithms 1-4 converges to stationary points in that region. This region can be viewed as an "incoherent neighborhood" of M, and can be formally defined as $K_1 \cap K_2 \cap K(\delta)$, where K_1, K_2 are defined as

$$\begin{aligned} K_1 &\triangleq \{(X, Y) | X \in \mathbb{R}^{m \times r}, Y \in \mathbb{R}^{n \times r}, \\ &\quad \|X^{(i)}\| \leq \beta_1, \|Y^{(j)}\| \leq \beta_2, \forall i, j\}, \\ K_2 &\triangleq \{(X, Y) | X \in \mathbb{R}^{m \times r}, Y \in \mathbb{R}^{n \times r}, \\ &\quad \|X\|_F \leq \beta_T, \|Y\|_F \leq \beta_T\}. \end{aligned} \tag{30}$$

and $K(\delta)$ is defined as

$$K(\delta) \triangleq \{(X, Y) | X \in \mathbb{R}^{m \times r}, Y \in \mathbb{R}^{n \times r}, \|M - XY^T\|_F \leq \delta\}. \tag{31}$$

Note that $K_2 = \Gamma(\beta_T)$ by our definition of Γ in (20). As mentioned in Section II-A, we only need to consider a Bernolli model of Ω where each entry is included into Ω with probability $p = \frac{S}{mn}$, where S satisfies (27).

The first lemma describes the local geometry and implies that any stationary point (X, Y) in $K_1 \cap K_2 \cap K(\delta)$ satisfies $XY^T = M$. The main steps to derive this geometrical property is described in Section I-C. The formal proof will be given in Section IV.

Lemma 3.1: There exist numerical constants C_0, C_d *such that the following holds. Assume* δ *is defined by* (16) *and* Ω *is generated by a Bernolli model with expected cardinality* S *satisfying* (27) *(i.e.* S *is lower bounded by the right hand side of* (27)*). Then, with probability at least* $1 - 1/n^4$, *the following holds: for all* $(X, Y) \in K_1 \cap K_2 \cap K(\delta)$, *there exist* $U \in \mathbb{R}^{m \times r}, \mathbb{V} \in R^{n \times r}$, *such that* $UV^T = M$ *and*

$$\begin{aligned} &\langle \nabla_X \tilde{F}(X, Y), X - U \rangle + \langle \nabla_Y \tilde{F}(X, Y), Y - V \rangle \\ &\quad \geq \frac{p}{4} \|M - XY^T\|_F^2. \end{aligned} \tag{32}$$

The second lemma describes the properties of the algorithms we presented. Throughout the paper, "under the same condition of Lemma 3.1" means "assume δ is defined by (16) and

Ω is generated by a Bernolli model with expected cardinality S satisfying (27), where C_0, C_d are the same numerical constants as those in Lemma 3.1". The proof of Lemma 3.2 will be given in Section V.

Lemma 3.2: Under the same conditions of Lemma 3.1, with probability at least $1-1/n^4$*, the sequence* (X_k, Y_k) *generated by either of Algorithms 1-4 has the following properties:*
(a) Each limit point of (X_k, Y_k) *is a stationary point of (P1).*
(b) $(X_k, Y_k) \in K_1 \cap K_2 \cap K(\delta)$, $\forall k \geq 0$.

Intuitively, $\|X_k^{(i)}\|, \|Y_k^{(j)}\|, \|X_k\|_F, \|Y_k\|_F$ are bounded because of the regularization terms we introduced and that the objective function is decreasing, and $\|M - X_k Y_k^T\|_F$ is bounded because the objective function is decreasing (however, the intuition is not enough and the proof requires some extra effort). In Section V we provide some easily verifiable conditions for Property (b) to hold (see Proposition 5.1), so that Lemma 3.2 and Theorem 3.1 can be extended to other algorithms.

With these two lemmas, the proof of Theorem 3.1 is quite straightforward and presented below.

Proof of Theorem 3.1: Consider any limit point (X_*, Y_*) of sequence $\{(X_k, Y_k)\}$ generated by either of Algorithms 1-4. According to Property (a) of Lemma (3.2), (X_*, Y_*) is a stationary point of problem (P1), i.e. $\nabla_X \tilde{F}(X_*, Y_*) = 0, \nabla_Y \tilde{F}(X_*, Y_*) = 0$. According to Property (b) of Lemma 3.2, with probability at least $1-1/n^4$, $(X_k, Y_k) \in K_1 \cap K_2 \cap K(\delta)$ for all k, implying $(X_*, Y_*) \in K_1 \cap K_2 \cap K(\delta)$. Then we can apply Lemma 3.1 by plugging $(X, Y) = (X^*, Y^*)$ into (32) to conclude that with probability at least $1-2/n^4$, $\|M - X_* Y_*^T\|_F \leq 0$, i.e. $X_* Y_*^T = M$. □

Remark: Note that $X_* Y_*^T = M$ does not necessarily imply the global optimality of (X_*, Y_*) since we have not proved $G(X_*, Y_*) = 0$. Nevertheless, the global optimality can be easily proved using a different version of Lemma 3.1 (see the discussion before Lemma 3.3); in other words, Theorem 3.1 can be slightly strengthened to "Algorithm 1-4 converge to the global optima of problem (P1)", instead of "Algorithm 1-4 recover *M*". The same argument can be used to show a more general result than Theorem 3.1, as stated in the following corollary.

Corollary 3.1: Under the same conditions of Theorem 3.1, any algorithm satisfying Properties (a) and (b) in Lemma 3.2 reconstructs M exactly with probability at least $1-2/n^4$.

B. Proof of Theorem 3.2

The proof of Theorem 3.2 applies a standard framework for first order methods: the convergence rate (or iteration complexity) can be derived from the "cost-to-go estimate" and the "sufficient descent" condition. For instance, the linear convergence $f(\boldsymbol{x}_k) - f^* \leq (1-c_1c_2)^k$ is a direct corollary of the cost-to-go estimate $\|\nabla f(\boldsymbol{x}_k)\|^2 \geq c_1[f(\boldsymbol{x}_k) - f^*]$ and the sufficient descent condition $f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1}) \geq c_2 \|\nabla f(\boldsymbol{x}_k)\|^2$, where f^* is the minimum value of f, and c_1, c_2 are certain constants. We remark that using other optimization frameworks may lead to stronger time complexity bounds; this is left as future work. For our problem, a variant of Lemma 3.1 can be viewed as the cost-to-go estimate; see Lemma 3.3 below. One difference with Lemma 3.1 is the following: for a stationary point (X_*, Y_*) that $\nabla \tilde{F}(X_*, Y_*) = 0$, Lemma 3.3 implies $\tilde{F}(X_*, Y_*) = 0$ (global optimality), but Lemma 3.1 implies $M = X_* Y_*^T$ (exact recovery). The relation between these two lemmas is that Lemma 3.3 is a direct consequence of (251), a slightly stronger version of Lemma 3.1. The main difficulties of proving the two lemmas are the same and lie in Proposition 4.1 and Proposition 4.2; see the formal proof in Appendix D.5.

Lemma 3.3 (Cost-To-Go Estimate): Under the same conditions of Lemma 3.1, with probability at least $1-1/n^4$*, the following holds:*

$$\|\nabla \tilde{F}(X, Y)\|_F^2 \geq \xi \tilde{F}(X, Y), \quad \forall (X, Y) \in K_1 \cap K_2 \cap K(\delta), \tag{33}$$

where $\xi = \frac{1}{C_g r^5 \kappa^3} p \Sigma_{\min}$ *(here* $C_g \geq 1$ *is a numerical constant).*

The following claim shows that Algorithm 1a satisfies the sufficient descent condition. It is easy to prove: it is well known that for minimizing a function (possibly nonconvex) with Lipschitz continuous gradient, the gradient descent method with constant step-size satisfies the sufficient decrease condition.

Claim 3.2 (Sufficient Descent): For the sequence $\boldsymbol{x}_k = (X_k, Y_k)$ *generated by Algorithm 1a (gradient descent with constant stepsize), we have*

$$\tilde{F}(\boldsymbol{x}_k) - \tilde{F}(\boldsymbol{x}_{k+1}) \geq \frac{\eta_1}{2} \|\nabla \tilde{F}(\boldsymbol{x}_k)\|_F^2, \tag{34}$$

where η_1 *is the stepsize bounded above by* $\bar{\eta}_1$ *defined in* (238).

The linear convergence can be easily derived from Lemma 3.1 and Claim 3.2. For completeness, we present the proof below.

Proof of Theorem 3.2: According to Property (b) of Lemma 3.2, with probability at least $1-1/n^4$, $(X_k, Y_k) \in K_1 \cap K_2 \cap K(\delta)$ for all k. According to Lemma 3.3 and Claim 3.2, we have (with probability at least $1-2/n^4$)

$$\tilde{F}(\boldsymbol{x}_k) - \tilde{F}(\boldsymbol{x}_{k+1}) \geq \frac{\eta_1}{2} \|\nabla \tilde{F}(\boldsymbol{x}_k)\|_F^2 \geq \frac{\eta_1}{2} \xi \tilde{F}(\boldsymbol{x}_k), \quad \forall k.$$

This relation can be rewritten as

$$\tilde{F}(\boldsymbol{x}_{k+1}) \leq (1 - \frac{1}{2}\eta_1 \xi) \tilde{F}(\boldsymbol{x}_k), \quad \forall k. \tag{35}$$

The stepsize η_1 can be bounded as $0 < \eta_1 \leq \bar{\eta}_1 \overset{(235)}{\leq} \frac{1}{4\beta_T^2} = \frac{1}{4C_T r \Sigma_{\max}} \leq \frac{1}{\Sigma_{\max}}$. Since $0 < \xi = \frac{1}{C_g r^5 \kappa^3} p \Sigma_{\min} \leq \Sigma_{\min}$, we have $0 < \eta_1 \xi \leq \frac{\Sigma_{\min}}{\Sigma_{\max}} \leq 1$, which implies $0 < 1 - \frac{1}{2}\eta_1 \xi < 1$. Then the relation (35) leads to

$$\tilde{F}(\boldsymbol{x}_k) \leq (1 - \frac{1}{2}\eta_1 \xi)^k \tilde{F}(\boldsymbol{x}_0), \quad \forall k,$$

which finishes the proof. □

The same argument can be used to show a more general result than Theorem 3.2, as stated in the following corollary.

Corollary 3.2: Under the same conditions of Theorem 3.1, any algorithm satisfying Properties (a) and (b) in Lemma 3.2 and the sufficient decrease condition (34) *has the linear convergence property, i.e. generates a sequence* (X_k, Y_k) *that satisfies* (28)*.*

IV. Proof of Lemma 3.1

In Section IV-A, we will show that to prove Lemma 3.1, we only need to construct U, V to satisfy three inequalities that $\|\mathcal{P}_\Omega((U-X)(V-Y)^T)\|_F$ and $\|((U-X)(V-Y)^T)\|_F$ are bounded above and $\langle \nabla_X G, X-U\rangle + \langle \nabla_Y G, Y-V\rangle$ is bounded below. In Section IV-B we describe two propositions that specify the choice of U, V, and then we show that such U, V satisfy the three desired inequalities in Section IV-B and subsequent subsections.

A. Preliminary Analysis

Since $(X, Y) \in K(\delta)$, we have

$$d \triangleq \|M - XY^T\|_F \le \delta \overset{(16)}{=} \frac{\Sigma_{\min}}{C_d r^{1.5}\kappa}. \tag{36}$$

To ensure (32) holds, we only need to ensure that the following two inequalities hold:

$$\phi_F = \langle \nabla_X F, X-U\rangle + \langle \nabla_Y F, Y-V\rangle \ge \frac{p}{4}d^2, \tag{37a}$$

$$\phi_G = \langle \nabla_X G, X-U\rangle + \langle \nabla_Y G, Y-V\rangle \ge 0. \tag{37b}$$

Define

$$a \triangleq U(Y-V)^T + (X-U)V^T, \quad b \triangleq (U-X)(V-Y)^T. \tag{38}$$

Then

$$XY^T - M = a + b, \quad (X-U)Y^T + X(Y-V)^T = a + 2b.$$

Using the expressions of $\nabla_X F, \nabla_Y F$ in (24), we bound ϕ_F as follows:

$$\begin{aligned}\phi_F &= \langle \nabla_X F, X-U\rangle + \langle \nabla_Y F, Y-V\rangle \\ &= \langle \mathcal{P}_\Omega(XY^T - M), (X-U)Y^T + X(Y-V)^T\rangle \\ &= \langle \mathcal{P}_\Omega(a+b), \mathcal{P}_\Omega(a+2b)\rangle \\ &= \|\mathcal{P}_\Omega(a)\|_F^2 + 2\|\mathcal{P}_\Omega(b)\|_F^2 + 3\langle \mathcal{P}_\Omega(a), \mathcal{P}_\Omega(b)\rangle \\ &\ge \|\mathcal{P}_\Omega(a)\|_F^2 + 2\|\mathcal{P}_\Omega(b)\|_F^2 - 3\|\mathcal{P}_\Omega(a)\|_F\|\mathcal{P}_\Omega(b)\|_F.\end{aligned} \tag{39}$$

The reason to decompose $M - XY^T$ as $a + b$ is the following. In order to bound $\|\mathcal{P}_\Omega(M - XY^T)\|_F$, we notice $E(\mathcal{P}_\Omega(M - XY^T)) = p(M - XY^T)$ and wish to prove $\|\mathcal{P}_\Omega(M - XY^T)\|_F^2 \approx O(pd^2)$. However, $\|\mathcal{P}_\Omega(A)\|_F$ could be as large as $\|A\|_F$ if the matrix A is not independent of the random subset Ω (e.g. choose A s.t. $A = \mathcal{P}_\Omega(A)$). This issue can be resolved by decomposing $XY^T - M$ as $a + b$ and bounding $\|\mathcal{P}_\Omega(a)\|_F$ and $\|\mathcal{P}_\Omega(b)\|_F$ separately. In fact, $\|\mathcal{P}_\Omega(a)\|_F$ can be bounded because a lies in a space spanned by the matrices with the same row space or column space as M, which is independent of Ω ([4, Th. 4.1]). $\|\mathcal{P}_\Omega(b)\|_F$ can be bounded according to a random graph lemma of [31], [46], which requires U, V, X, Y to be incoherent (i.e. have bounded row norm).

We claim that (37a) is implied by the following two inequalities:

$$\|\mathcal{P}_\Omega(b)\|_F = \|\mathcal{P}_\Omega((U-X)(V-Y)^T)\|_F \le \frac{1}{5}\sqrt{p}d; \tag{40a}$$

$$\|b\|_F = \|(U-X)(V-Y)^T\|_F \le \frac{1}{10}d. \tag{40b}$$

In fact, assume (40a) and (40b) are true, we prove $\phi_F \ge pd^2/4$ as follows. By $XY^T - M = a + b$ we have

$$\|a\|_F \ge \|M - XY^T\|_F - \|b\|_F \overset{(40b)}{\ge} \frac{9}{10}d. \tag{41}$$

Recall that the SVD of M is $M = \hat{U}\Sigma\hat{V}^T$ and M satisfies the incoherence condtion (12). It follows from $M = UV^T = \hat{U}\Sigma\hat{V}^T$ that $M, U, \hat{U}$ have the same column space, thus there exists some matrix $B_1 \in \mathbb{R}^{r\times r}$ such that $U = \hat{U}B_1$; similarly, there exists $B_2 \in \mathbb{R}^{r\times r}$ such that $V = \hat{V}B_2$. Therefore, by the definition of a in (38) we have

$$a \in \mathcal{T} \triangleq \{\hat{U}W_2^T + W_1\hat{V}^T \mid W_1 \in \mathbb{R}^{m\times r}, W_2 \in \mathbb{R}^{n\times r}\}. \tag{42}$$

By [4, Theorem 4.1], for $|\Omega|$ satisfying (27) with large enough C_0, we have that with probability at least $1 - 1/(2n^4)$, $\|\mathcal{P}_\mathcal{T}\mathcal{P}_\Omega\mathcal{P}_\mathcal{T}(a) - p\mathcal{P}_\mathcal{T}(a)\|_F \le \frac{1}{6}p\|a\|_F$ (note that this bound holds uniformly for all $a \in \mathcal{T}$, thus also holds when a is dependent on Ω). Since $a \in \mathcal{T}$, this inequality can be simplified to

$$\|\mathcal{P}_\mathcal{T}\mathcal{P}_\Omega(a) - pa\|_F \le \frac{1}{6}p\|a\|_F. \tag{43}$$

Following the analysis of [4, Corollary 4.3], we have

$$\begin{aligned}\|\mathcal{P}_\Omega(a)\|_F^2 &= \|\mathcal{P}_\Omega\mathcal{P}_\mathcal{T}(a)\|_F^2 \\ &= \langle a, \mathcal{P}_\mathcal{T}\mathcal{P}_\Omega^2\mathcal{P}_\mathcal{T}(a)\rangle = \langle a, \mathcal{P}_\mathcal{T}\mathcal{P}_\Omega(a)\rangle \\ &= \langle a, pa\rangle + \langle a, \mathcal{P}_\mathcal{T}\mathcal{P}_\Omega(a) - pa\rangle.\end{aligned} \tag{44}$$

The absolute value of the second term can be bounded as

$$|\langle a, \mathcal{P}_\mathcal{T}\mathcal{P}_\Omega(a) - pa\rangle| \le \|a\|_F\|\mathcal{P}_\mathcal{T}\mathcal{P}_\Omega(a) - pa\|_F \overset{(43)}{\le} \frac{1}{6}p\|a\|_F^2,$$

which implies $-\frac{1}{6}p\|a\|_F^2 \le \langle a, \mathcal{P}_\mathcal{T}\mathcal{P}_\Omega(a) - pa\rangle \le \frac{1}{6}p\|a\|_F^2$. Substituting into (44), we obtain that with probability at least $1 - 1/(2n^4)$,

$$\frac{5}{6}\|a\|_F^2 \le \|\mathcal{P}_\Omega(a)\|_F^2 \le \frac{7}{6}\|a\|_F^2. \tag{45}$$

The first inequality of the above relation implies

$$\|\mathcal{P}_\Omega(a)\|_F^2 \ge \frac{5}{6}\|a\|_F^2 \overset{(41)}{\ge} \frac{27}{40}pd^2. \tag{46}$$

According to (39) and the bounds (46) and (40a), we have $\phi_F/(pd^2) \ge \frac{27}{40} + 2(\frac{1}{5})^2 - \frac{3}{5}\sqrt{\frac{27}{40}} \ge \frac{1}{4}$, which proves (37a).

In summary, to find a factorization $M = UV^T$ such that (32) holds, we only need to ensure that the factorization satisfies (40b), (40a) and (37b). In the following three subsections, we will show that such a factorization $M = UV^T$ exists. Specifically, U, V will be defined in Table VII and the three desired inequalities will be proved in Corollary 4.2, Proposition 4.3 and Claim 4.1 respectively.

B. Definitions of U, V and Key Technical Results

We construct U, V according to two propositions, which will be stated in this subsection and proved in the appendix. The first proposition states that if XY^T is close to M, then there exists a factorization $M = UV^T$ such that U (resp. V) is close to X (resp. Y), and U, V are incoherent. Roughly speaking, this proposition shows the continuity of the factorization map $Z = XY^T \mapsto (X, Y)$ near a low-rank matrix M. The condition $X, Y \in K_1 \cap K_2 \cap K(\delta)$ and (16) implies that $d \triangleq \|M - XY^T\|_F \le \delta = \frac{\Sigma_{\min}}{C_d r^{1.5}\kappa}$ and $\|X\|_F \le \beta_T, \|Y\|_F \le \beta_T$, thus for large enough C_d, the assumptions of Proposition 4.1 hold. Similarly, the assumptions of the other results in this subsection also hold.

Proposition 4.1: Suppose $M \in \mathbb{R}^{m\times n}$ is a rank-r matrix with $\Sigma_{\max}$ ($\Sigma_{\min}$) being the largest (smallest) non-zero singular value, and M is μ-incoherent. There exists a numerical constant C_T such that the following holds: If

$$d \triangleq \|M - XY^T\|_F \le \frac{\Sigma_{\min}}{11r}, \tag{47a}$$

$$\|X\|_F \le \beta_T, \quad \|Y\|_F \le \beta_T, \tag{47b}$$

where $\beta_T = \sqrt{C_T r \Sigma_{\max}}$, then there exist $U \in \mathbb{R}^{m\times r}$, $V \in \mathbb{R}^{n\times r}$ such that

$$UV^T = M, \tag{48a}$$

$$\|U\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|X\|_F, \tag{48b}$$

$$\|U - X\|_F \le \frac{6\beta_T}{5\Sigma_{\min}}d, \quad \|V - Y\|_F \le \frac{3\beta_T}{\Sigma_{\min}}d, \tag{48c}$$

$$\|U^{(i)}\|^2 \le \frac{r\mu}{m}\beta_T^2, \quad \|V^{(j)}\|^2 \le \frac{3r\mu}{2n}\beta_T^2. \tag{48d}$$

The proof of Proposition 4.1 is given in Appendix A.2.

Remark 1: A symmetric result that switches X, U and Y, V in the above proposition holds: under the conditions of Proposition (4.1), there exist U, V satisfying (48) with U, V reversed, i.e. $UV^T = M$, $\|V\|_F(1 - \frac{d}{\Sigma_{\min}}) \le \|Y\|_F$, $\|U - X\|_F \le \frac{3\beta_T}{\Sigma_{\min}}d, \|V - Y\|_F \le \frac{6\beta_T}{5\Sigma_{\min}}d$, and $\|U^{(i)}\|^2 \le \frac{3r\mu}{2m}\beta_T^2, \|V^{(j)}\|^2 \le \frac{r\mu}{n}\beta_T^2$.

Remark 2: To prove Theorem 3.1 (convergence), we only need $\|U\|_F \le \|X\|_F$; here the slightly stronger requirement $\|U\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|X\|_F$ is for the purpose of proving Theorem 3.2 (linear convergence).

Remark 3: Without the incoherence assumption on M, by the same proof we can show that there still exist U, V satisfying (48a) and (48c), i.e. $M = UV^T$ and U, V are close to X, Y respectively. Such a result bears some similarity with the classical perturbation theory for singular value decomposition [45]. In particular, [45] proved that for two low-rank matrices[4] that are close, the spaces spanned by the left (resp. right) singular vectors of the two matrices are also close. Note that the singular vectors themselves may be very sensitive to perturbations and no such perturbation bounds can be established (see [63, Sec. 6]). The difference of our work with the classical perturbation theory is that we do not consider SVD of two matrices; instead, we allow one matrix to have an arbitrary factorization, and the factorization of the other matrix can be chosen accordingly. Since we do not have any restriction on the factorization XY^T (except the dimensions) and the norms of X and Y can be arbitrarily large, the distance between two corresponding factors has to be proportional to the norm of one single factor, which explains the coefficient β_T in (48c).

[4]The result in [45] also covered the case of two approximately low-rank matrices, but we only consider the case of exact low-rank matrices here.

Unfortunately, Proposition 4.1 is not strong enough to prove $\phi_G \ge 0$ when both $\|X\|_F$ and $\|Y\|_F$ are large (see an analysis in Section IV-D). To resolve this issue, we need to prove the second proposition in which there is an additional assumption that both $\|X\|_F$ and $\|Y\|_F$ are large, and an additional requirement that both $\|U\|_F$ and $\|V\|_F$ are bounded (by the norms of original factors $\|X\|_F$ and $\|Y\|_F$ respectively). More specifically, the proposition states that if M is close to XY^T, and both $\|X\|_F$ and $\|Y\|_F$ are large, then there is a factorization $M = UV^T$ such that U (resp. V) is close to X (resp. Y), and $\|U\|_F \le \|X\|_F, \|V\|_F \le \|Y\|_F$. For the purpose of proving linear convergence, we prove a slightly stronger result that $\|V\|_F \le (1-d/\Sigma_{\min})\|Y\|_F$. The previous result Proposition 4.1 can be viewed as a perturbation analysis for an arbitrary factorization, while Proposition 4.2 can be viewed as an enhanced perturbation analysis for a constrained factorization. Although Proposition 4.2 is just a simple variant of Proposition 4.1, it seems to require a much more involved proof than Proposition 4.1. See the formal proof of Proposition 4.2 in Appendix B.1.

Proposition 4.2: Suppose $M \in \mathbb{R}^{m\times n}$ is a rank-r matrix with $\Sigma_{\max}$ ($\Sigma_{\min}$) being the largest (smallest) non-zero singular value, and M is μ-incoherent. There exist numerical constants C_d, C_T such that the following holds: if

$$d \triangleq \|M - XY^T\|_F \le \frac{\Sigma_{\min}}{C_d r}, \tag{49a}$$

$$\sqrt{\frac{2}{3}}\beta_T \le \|X\|_F \le \beta_T, \quad \sqrt{\frac{2}{3}}\beta_T \le \|Y\|_F \le \beta_T, \tag{49b}$$

where $\beta_T = \sqrt{C_T r \Sigma_{\max}}$, then there exist $U \in \mathbb{R}^{m\times r}$, $V \in \mathbb{R}^{n\times r}$ such that

$$UV^T = M, \tag{50a}$$

$$\|U\|_F \le \|X\|_F, \quad \|V\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F, \tag{50b}$$

$$\|U - X\|_F\|V - Y\|_F \le 65\sqrt{r}\frac{\beta_T^2}{\Sigma_{\min}^2}d^2,$$

$$\max\{\|U - X\|_F, \|V - Y\|_F\} \le \frac{17}{2}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d, \tag{50c}$$

$$\|U^{(i)}\|^2 \le \frac{r\mu}{m}\beta_T^2, \quad \|V^{(j)}\|^2 \le \frac{r\mu}{n}\beta_T^2. \tag{50d}$$

Remark: A symmetric result that switches X, U and Y, V in the above proposition still holds; the only change is that (50b) will become $\|U\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|X\|_F, \ \|V\|_F \le \|Y\|_F$. It is easy to prove a variant of the above proposition in which (50b) is changed to $\|U\|_F \le (1 - \frac{d}{2\Sigma_{\min}})\|X\|_F, \|V\|_F \le (1 - \frac{d}{2\Sigma_{\min}})\|Y\|_F$; in other words, the asymmetry of X, U and Y, V in (50b) is artificial. Nevertheless, Proposition 4.2 is enough for our purpose.

Throughout the proof of Lemma 3.1, U, V are defined in Table IV-B.

According to Proposition 4.1 and Proposition 4.2 (and their symmetric results), the properties of U, V defined in Tabel VII are summarized in the following corollary. For simplicity, we only present the case that $\|X\|_F \le \|Y\|_F$; in the other case that $\|X\|_F > \|Y\|_F$, a symmetric result of Corollary 4.1 holds.

TABLE VII

DEFINITION OF U, V

Definition of U, V in different cases
Case 1: $\|X\|_F \le \|Y\|_F$.
Case 1.1 : $\|X\|_F < \sqrt{\frac{2}{3}}\beta_T$. Define U, V according to the symmetrical result of Proposition IV.1, i.e. U, V satisfy (48) with X, U and Y, V reversed.
Case 1.2: $\|X\|_F, \|Y\|_F \in [\sqrt{\frac{2}{3}}\beta_T, \beta_T]$. Define U, V according to Proposition IV.2.
Case 2: $\|Y\|_F < \|X\|_F$.
Similar to Case 1 but with the roles of X, U and Y, V reversed.

Corollary 4.1: Suppose $d \triangleq \|XY^T - M\|_F \le \frac{\Sigma_{\min}}{C_d r}$ and $\|X\|_F \le \|Y\|_F$, then U, V defined in Table VII satisfy:

$$UV^T = M; \tag{51a}$$

$$\|U - X\|_F \|V - Y\|_F \le 65\sqrt{r}\frac{\beta_T^2}{\Sigma_{\min}^2}d^2;$$

$$\max\{\|U - X\|_F, \|V - Y\|_F\} \le \frac{17}{2}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d, \tag{51b}$$

$$\|U^{(i)}\|^2 \le \frac{3}{2}\frac{r\mu}{m}\beta_T^2, \quad \|V^{(j)}\|^2 \le \frac{3}{2}\frac{r\mu}{n}\beta_T^2; \tag{51c}$$

$$\|V\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F;$$

$$\text{if } \|X\|_F > \sqrt{\frac{2}{3}}\beta_T, \text{ then } \|U\|_F \le \|X\|_F. \tag{51d}$$

In (51b), we bound $\|U - X\|_F \|V - Y\|_F$ by $O(d^2)$ with a rather complicated coefficient, but to prove (40b) we need a bound $O(d)$ with a coefficient $1/10$. Under a slightly stronger condition on d than that of Corollary 4.1, which still holds for $(X, Y) \in K(\delta)$ with δ defined in (16), we can prove the bound (40b) by (51b).

Corollary 4.2: There exists a numerical constant C_d such that if

$$d \triangleq \|M - XY^T\|_F \le \frac{\Sigma_{\min}}{C_d r^{1.5}\kappa}, \tag{52}$$

then U, V defined in Table VII satisfy (40b).

Proof of Corollary 4.2: According to (51b) , we have

$$\begin{aligned}\|U - X\|_F \|V - Y\|_F &\le 65\frac{\beta_T^2}{\Sigma_{\min}^2}\sqrt{r}d^2 = 65C_T r^{1.5}\frac{\Sigma_{\max}}{\Sigma_{\min}^2}d^2 \\ &= 65C_T r^{1.5}\kappa\frac{d}{\Sigma_{\min}}d \le \frac{1}{10}d,\end{aligned}$$

where the last inequliaty follows from (52) with $C_d \ge 650C_T$. □

In the next two subsections, we will use the properties in Corollary 4.1 to prove (40a) and (37b).

C. Upper Bound on $\|\mathcal{P}_\Omega((U - X)(V - Y)^T)\|_F$

The following result states that for U, V defined in Table VII, (40a) holds.

Proposition 4.3: Under the same conditions as Lemma 3.1, with probability at least $1 - 1/(2n^4)$, the following is true. For any $(X, Y) \in K_1 \cap K_2 \cap K(\delta)$ and U, V defined in Table VII, we have

$$\|\mathcal{P}_\Omega((U - X)(V - Y)^T)\|_F^2 \le \frac{p}{25}\|M - XY^T\|_F^2. \tag{53}$$

Proof of Proposition 4.3: We need the following random graph lemma [31, Lemma 7.1].

Lemma 4.1: There exist numerical constants C_0, C_1 such that if $|\Omega| \ge C_0\sqrt{\alpha}n\log n$, then with probability at least $1 - 1/(2n^4)$, *for all $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$,*

$$\sum_{(i,j)\in\Omega} x_i y_j \le C_1 p\|x\|_1\|y\|_1 + C_1\alpha^{\frac{3}{4}}\sqrt{np}\|x\|_2\|y\|_2. \tag{54}$$

Let $Z = U - X, W = V - Y$ and $z_i = \|Z^{(i)}\|^2$, $w_j = \|W^{(j)}\|^2$. We have

$$\begin{aligned}\|\mathcal{P}_\Omega((U - X)(V - Y)^T)\|_F^2 &= \sum_{(i,j)\in\Omega}(ZW^T)_{ij}^2 \\ &\le \sum_{(i,j)\in\Omega}\|Z^{(i)}\|^2\|W^{(j)}\|^2 = \sum_{(i,j)\in\Omega} z_i w_j.\end{aligned} \tag{55}$$

Invoking Lemma 4.1, we have

$$\begin{aligned}&\|\mathcal{P}_\Omega((U - X)(V - Y)^T)\|_F^2 \\ &\quad\le C_1 p\|z\|_1\|w\|_1 + C_1\alpha^{\frac{3}{4}}\sqrt{np}\|z\|_2\|w\|_2.\end{aligned} \tag{56}$$

Analogous to the proof of (40b) in Corollary 4.2, we can prove that $\|U - X\|_F\|V - Y\|_F \le d/(10\sqrt{C_1})$ for large enough C_d (in fact, $C_d \ge 650C_T\sqrt{C_1}$ suffices). Therefore, we have

$$\begin{aligned}\|z\|_1\|w\|_1 &= \|Z\|_F^2\|W\|_F^2 = \|U - X\|_F^2\|V - Y\|_F^2 \\ &\le \frac{1}{100C_1}d^2.\end{aligned} \tag{57}$$

We still need to bound $\|z\|_2$ and $\|w\|_2$. We have

$$\begin{aligned}\|z\|_2 &= \sqrt{\sum_i \|Z^{(i)}\|^4} \\ &\le \sqrt{\max_i \|Z^{(i)}\|^2 \sum_j \|Z^{(j)}\|^2} \\ &\le \max_i(\|U^{(i)}\| + \|X^{(i)}\|)\|U - X\|_F \\ &\le (\sqrt{\frac{3r\mu}{2m}}\beta_T + \beta_1)\|U - X\|_F \\ &\le \sqrt{8}\sqrt{\frac{r\mu}{m}}\beta_T\|U - X\|_F.\end{aligned} \tag{58}$$

Here, the third inequliaty follows from the property (51c) in Corollary 4.1 and the condition $(X, Y) \in K_1$ (which implies $\|X^{(i)}\| \le \beta_1$), and the fourth inequliaty follows from the definition of β_1 in (15). Similarly,

$$\|w\|_2 \le \max_j(\|V^{(j)}\| + \|Y^{(j)}\|)\|V - Y\|_F \le \sqrt{8}\sqrt{\frac{r\mu}{n}}\beta_T\|V - Y\|_F. \quad (59)$$

Multiplying (58) and (59), we get

$$\begin{aligned} \|z\|_2\|w\|_2 &\le 8\frac{r\mu}{\sqrt{mn}}\beta_T^2\|U - X\|_F\|V - Y\|_F \\ &\overset{(51b)}{\le} 8\frac{r\mu}{\sqrt{mn}}\beta_T^2 65\sqrt{r}\frac{\beta_T^2}{\Sigma_{\min}^2}d^2 \\ &\overset{(15)}{=} 520C_T^2\frac{1}{\sqrt{mn}}\mu r^{3.5}\kappa^2 d^2. \end{aligned}$$

Thus the second term in (56) can be bounded as

$$\begin{aligned} C_1\alpha^{\frac{3}{4}}\sqrt{np}\|z\|_2\|w\|_2 &\le 520C_1C_T^2\frac{\alpha^{\frac{3}{4}}\sqrt{np}}{\sqrt{mn}}\mu r^{3.5}\kappa^2 d^2 \\ &\le \frac{3}{100}pd^2, \end{aligned} \quad (60)$$

where the last inequality is equivalent to $520^2C_1^2C_T^4\alpha^{\frac{3}{2}}\mu^2r^7\kappa^4 \le \frac{9}{100^2}|\Omega|/n$, which holds due to (27) with large enough numerical constant C_0. Plugging (57) and (60) into (56), we get $\|\mathcal{P}_\Omega((U - X)(V - Y)^T)\|_F^2 \le \frac{p}{25}d^2 = \frac{p}{25}\|M - XY^T\|_F^2$. □

D. Lower Bound on ϕ_G

In this subsection, we prove the following claim.

Claim 4.1: U, V defined in Table VII satisfy (37b), i.e. $\phi_G = \langle\nabla_X G, X - U\rangle + \langle\nabla_Y G, Y - V\rangle \ge 0$.

Proof of Claim 4.1: By the expressions of $\nabla_X G, \nabla_Y G$ in (24), we have

$$\begin{aligned} \phi_G &= \langle\nabla_X G, X - U\rangle + \langle\nabla_Y G, Y - V\rangle \\ &= \rho\sum_{i=1}^m G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2})\frac{3}{\beta_1^2}\langle X^{(i)}, X^{(i)} - U^{(i)}\rangle \\ &\quad +\rho G_0'(\frac{3\|X\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle X, X - U\rangle \\ &\quad +\rho\sum_{j=1}^n G_0'(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2})\frac{3}{\beta_2^2}\langle Y^{(j)}, Y^{(j)} - V^{(j)}\rangle \\ &\quad +\rho G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle Y, Y - V\rangle, \end{aligned} \quad (61)$$

where $G_0'(z) = I_{[1,\infty]}(z)2(z - 1)$.

Firstly, we prove

$$h_{1i} \triangleq G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2})\frac{3}{\beta_1^2}\langle X^{(i)}, X^{(i)} - U^{(i)}\rangle \ge 0, \quad \forall\ i, \quad (62a)$$

$$h_{3j} \triangleq G_0'(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2})\frac{3}{\beta_2^2}\langle Y^{(j)}, Y^{(j)} - V^{(j)}\rangle \ge 0, \quad \forall\ j. \quad (62b)$$

We only need to prove (62a); the proof of (62b) is similar. We consider two cases.

Case 1: $\|X^{(i)}\|^2 \le \frac{2\beta_1^2}{3}$. Note that $\frac{3\|X^{(i)}\|^2}{2\beta_1^2} \le 1$ implies $G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}) = 0$, thus $h_{1i} = 0$.

Case 2: $\|X^{(i)}\|^2 > \frac{2\beta_1^2}{3}$. By Corollary 4.1 and the fact that $\beta_1^2 = \beta_T^2\frac{3\mu r}{m}$, we have

$$\|U^{(i)}\|^2 \le \frac{3r\mu}{2m}\beta_T^2 \le \frac{2\beta_1^2}{3} < \|X^{(i)}\|^2. \quad (63)$$

As a result, $\langle X^{(i)}, X^{(i)}\rangle = \|X^{(i)}\|\|X^{(i)}\| > \|X^{(i)}\|\|U^{(i)}\| \ge \langle X^{(i)}, U^{(i)}\rangle$, which implies $\langle X^{(i)}, X^{(i)} - U^{(i)}\rangle \ge 0$. Combining this inequality with the fact that $G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}) \ge 0$, we get $h_{1i} \ge 0$.

Secondly, we prove

$$h_2 + h_4 \ge 0,$$

where

$$\begin{aligned} h_2 &\triangleq G_0'(\frac{3\|X\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle X, X - U\rangle, \\ h_4 &\triangleq G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle Y, Y - V\rangle. \end{aligned} \quad (64)$$

Without loss of generality, we can assume $\|X\|_F \le \|Y\|_F$, and we will apply Corollary 4.1 to prove (64). If $\|Y\|_F < \|X\|_F$, we can apply a symmetric result of Corollary 4.1 to prove (64). We further consider three cases.

Case 1: $\|X\|_F \le \|Y\|_F \le \sqrt{\frac{2}{3}}\beta_T$. In this case $G_0'(\frac{3\|X\|_F^2}{2\beta_T^2}) = G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2}) = 0$, which implies $h_2 = h_4 = 0$, thus (64) holds.

Case 2: $\|X\|_F \le \sqrt{\frac{2}{3}}\beta_T < \|Y\|_F$. Then $G_0'(\frac{3\|X\|_F^2}{2\beta_T^2}) = 0$, which implies $h_2 = 0$. By (51d) in Corollary 4.1 we have $\|V\|_F \le \|Y\|_F$, which implies $\langle Y, Y\rangle \ge \|Y\|_F\|V\|_F \ge \langle Y, V\rangle$, i.e. $\langle Y, Y - V\rangle \ge 0$. Combined with the nonnegativity of $G_0'(\cdot)$, we get $h_4 \ge 0$. Thus $h_2 + h_4 = h_4 \ge 0$.

Case 3: $\sqrt{\frac{2}{3}}\beta_T < \|X\|_F \le \|Y\|_F$. By (51d) in Corollary 4.1, we have $\|U\|_F \le \|X\|_F$ and $\|V\|_F \le \|Y\|_F$. Similar to the argument in Case 2 we can prove $h_2 \ge 0, h_4 \ge 0$ and (64) follows.

In all three cases, we have proved (64), thus (64) holds.

We conclude that for U, V defined in Table VII,

$$\phi_G \overset{(61)}{=} \rho\left(\sum_i h_{1i} + \sum_j h_{3j} + h_2 + h_4\right) \overset{(62),(64)}{\ge} 0,$$

which finishes the proof of Claim 4.1. □

Remark: Based on the above proof, we can explain why Proposition 4.1 is not enough to prove $\phi_G \ge 0$. Note that $h_2 = 0$ when $\|X\|_F > \sqrt{\frac{2}{3}}\beta_T$ and $h_4 = 0$ when $\|Y\|_F > \sqrt{\frac{2}{3}}\beta_T$. To prove $h_2 \ge 0, h_4 \ge 0$, it suffices to prove: (i) $\|U\|_F \le \|X\|_F$ when $\|X\|_F > \sqrt{\frac{2}{3}}\beta_T$; (ii) $\|V\|_F \le \|Y\|_F$ when $\|Y\|_F > \sqrt{\frac{2}{3}}\beta_T$. For the choice of U, V in Proposition 4.1, we have $\|U\|_F \le \|X\|_F$, but there is no guarantee that (ii) holds. Similarly, for the choice of U, V in the symmetric result of Proposition 4.1, we have $\|V\|_F \le \|Y\|_F$, but there is no guarantee that (i) holds. Thus, Proposition 4.1

is not enough to prove $\phi_G \geq 0$. To guarantee that (i) and (ii) hold simultaneously, we need a complementary result for the case $\|X\|_F > \sqrt{\frac{2}{3}}\beta_T, \|Y\|_F > \sqrt{\frac{2}{3}}\beta_T$. This motivates our Proposition 4.2.

V. Proof of Lemma 3.2

Property (a) in Lemma 3.2 (convergence to stationary points) is a basic requirement for many reasonable algorithms and can be proved using classical results in optimization, so the difficulty mainly lies in how to prove Property (b). We will give some easily verifiable conditions for Property (b) to hold and then show that Algorithms 1-4 satisfy these conditions. This proof framework can be used to extend Theorem 3.1 to many other algorithms.

The following claim states that Algorithms 1-4 satisfy Property (a). The proof of this claim is given in Appendix D.5.

Claim 5.1: Suppose Ω satisfies (29)*, then each limit point of the sequence generated by Algorithms 1-4 is a stationary point of problem (P1).*

For Property (b), we first show that the initial point (X_0, Y_0) lies in an incoherent neighborhood $(\sqrt{\frac{2}{3}}K_1) \cap (\sqrt{\frac{2}{3}}K_2) \cap K_{\delta_0}$, where cK_i denotes the set $\{(cX, cY) \mid (X, Y) \in K_i\}, i = 1, 2$. The proof of Claim 5.2 will be given in Appendix D.1. The purpose of proving $(X_0, Y_0) \in (\sqrt{\frac{2}{3}}K_1) \cap (\sqrt{\frac{2}{3}}K_2)$ rather than $(X_0, Y_0) \in K_1 \cap K_2$ is to guarantee that $G(X_0, Y_0) = 0$, where G is the regularizer defined in (13).

Claim 5.2: Under the same condition of Lemma 3.1, with probability at least $1 - 1/(2n^4)$, (X_0, Y_0) given by the procedure Initialize *belongs to $(\sqrt{\frac{2}{3}}K_1) \cap (\sqrt{\frac{2}{3}}K_2) \cap K_{\delta_0}$, where δ_0 is defined by* (16)*, i.e.*

(a) $\|X_0^{(i)}\| \leq \sqrt{\frac{2}{3}}\beta_1, i = 1, 2, \ldots, m;\quad \|Y_0^{(j)}\| \leq \sqrt{\frac{2}{3}}\beta_2, j = 1, \ldots, n;$

(b) $\|X_0\|_F \leq \sqrt{\frac{2}{3}}\beta_T,\ \|Y_0\|_F \leq \sqrt{\frac{2}{3}}\beta_T;$

(c) $\|M - X_0Y_0^T\|_F \leq \delta_0.$

The next result provides some general conditions for (X_t, Y_t) to lie in $K_1 \cap K_2 \cap K(\delta)$. To simplify the notations, denote $\boldsymbol{x}_t \triangleq (X_t, Y_t)$ and

$$\boldsymbol{u}^* \triangleq (\hat{U}\Sigma^{1/2}, \hat{V}\Sigma^{1/2}),$$

where $\hat{U}\Sigma\hat{V}$ is the SVD of M. Recall that $\tilde{F}(\boldsymbol{u}^*) = 0$ (proved in the paragraph after (19)). We say a function $\psi(\bar{\boldsymbol{x}}, \boldsymbol{\Delta}; \lambda)$ is a convex tight upper bound of $\tilde{F}(\boldsymbol{x})$ along the direction $\boldsymbol{\Delta}$ at $\bar{\boldsymbol{x}}$ if

$$\psi(\bar{\boldsymbol{x}}, \boldsymbol{\Delta}; \lambda) \text{ is convex over } \lambda \in \mathbb{R}; \tag{65a}$$

$$\psi(\bar{\boldsymbol{x}}, \boldsymbol{\Delta}; \lambda) \geq \tilde{F}(\bar{\boldsymbol{x}} + \lambda\boldsymbol{\Delta}),\ \forall\, \lambda \in \mathbb{R};\quad \psi(\bar{\boldsymbol{x}}, \boldsymbol{\Delta}; 0) = \tilde{F}(\bar{\boldsymbol{x}}). \tag{65b}$$

For example, $\psi(\bar{\boldsymbol{x}}, \boldsymbol{\Delta}; \lambda) = \tilde{F}(\bar{\boldsymbol{x}} + \lambda\boldsymbol{\Delta})$ satisfies (65) for either $\boldsymbol{\Delta} = (X, 0)$ or $\boldsymbol{\Delta} = (0, Y)$, where $X \in \mathbb{R}^{m\times r}$ and $Y \in \mathbb{R}^{n\times r}$ are arbitrary matrices. This definition is motivated by the block successive upper bound minimization method [55].

The proof of Proposition 5.1 is given in Appendix A.2.

Proposition 5.1: Suppose the sample set Ω satisfies (29) *and δ, δ_0 are defined by* (16)*. Consider an algorithm that starts from a point $\boldsymbol{x}_0 = (X_0, Y_0)$ and generates a sequence $\{\boldsymbol{x}_t\} = \{(X_t, Y_t)\}$. Suppose $\boldsymbol{x}_0$ satisfies*

$$\boldsymbol{x}_0 \in (\sqrt{\frac{2}{3}}K_1) \cap (\sqrt{\frac{2}{3}}K_2) \cap K(\delta_0), \tag{66}$$

and $\{\boldsymbol{x}_t\}$ satisfies either of the following three conditions:

$$1)\quad \tilde{F}(\boldsymbol{x}_t + \lambda\boldsymbol{\Delta}_t) \leq 2\tilde{F}(\boldsymbol{x}_0), \forall\ \lambda \in [0, 1], \quad \textit{where } \boldsymbol{\Delta}_t = \boldsymbol{x}_{t+1} - \boldsymbol{x}_t,\ \ \forall\ t; \tag{67a}$$

$$2)\quad 1 = \arg\min_{\lambda\in\mathbb{R}} \psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; \lambda), \quad \textit{where } \psi \textit{ satisfies } (65), \boldsymbol{\Delta}_t = \boldsymbol{x}_{t+1} - \boldsymbol{x}_t,\ \ \forall\ t; \tag{67b}$$

$$3)\quad \tilde{F}(\boldsymbol{x}_t) \leq 2\tilde{F}(\boldsymbol{x}_0),\quad d(\boldsymbol{x}_t, \boldsymbol{x}_0) \leq \frac{5}{6}\delta,\ \forall\ t. \tag{67c}$$

Then $\boldsymbol{x}_t = (X_t, Y_t) \in K_1 \cap K_2 \cap K(2\delta/3)$, for all $t \geq 0$.

The first condition means that $\tilde{F}$ is bounded above by $2\tilde{F}(\boldsymbol{x}_0)$ over the line segment between $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ for any t. This condition holds for gradient descent or SGD with small enough stepsize (see Claim 5.3). The second condition means that the new point $\boldsymbol{x}_{t+1}$ is the minimum of a convex tight upper bound of the original function along the direction $\boldsymbol{x}_{t+1} - \boldsymbol{x}_t$, and holds for BCD type methods such as Algorithm 2 and Algorithm 3 (see Claim 5.3). Note that the gradient descent method with exact line search stepsize does not satisfy this condition since $\tilde{F}$ is not jointly convex in the variable (X, Y). The third condition means that $\tilde{F}(\boldsymbol{x}_t)$ is bounded above and $\boldsymbol{x}_t$ is not far from $\boldsymbol{x}_0$ for any t. For standard nonlinear optimization algorithms, it is not easy to prove that $\boldsymbol{x}_t$ is not far from $\boldsymbol{x}_0$. However, as done by Algorithm 1 with restricted Armijo rule or restricted line search, we can force $d(\boldsymbol{x}_t, \boldsymbol{x}_0) \leq \frac{5}{6}\delta$ to hold when computing the new point $\boldsymbol{x}_t$.

The following claim shows that each of Algorithm 1-4 satisfies one of the three conditions in (67). The proof of Claim 5.3 is given in Appendix D.4.

Claim 5.3: The sequence $\{\boldsymbol{x}_t\}$ generated by Algorithm 1 with either restricted Armijo rule or restricted line search satisfies (67c)*. The sequence $\{\boldsymbol{x}_t\}$ generated by either Algorithm 2 or Algorithm 3 satisfies* (67b)*. Suppose the sample set Ω satisfies* (29)*, then the sequence $\{\boldsymbol{x}_t\}$ generated by either Algorithm 1 with constant stepsize or Algorithm 4 satisfies* (67a)*.*

To put things together, Claim 5.1 shows Algorithms 1-4 satisfy Property (a), and Proposition 5.1 together with Claim 5.2 and Claim 5.3 shows that Algorithms 1-4 satisfy Property (b). Therefore, we have proved Lemma 3.2.

Appendix

A. Supplemental Material for Section 2

A.1 Proof of Claim 2.1

This proof is quite straightforward and we mainly use the triangular inequalities and the boundedness of the considered region $\Gamma(\beta_0)$. In this proof, $f'(x)$ denotes the derivative of a function f at x.

Since $(X, Y), (U, V)$ belong to $\Gamma(\beta_0)$, we have

$$\begin{aligned} \|X\|_F \leq \beta_0, \quad \|Y\|_F \leq \beta_0, \\ \|U\|_F \leq \beta_0, \quad \|V\|_F \leq \beta_0. \end{aligned} \tag{68}$$

We first prove

$$\|\nabla F(X,Y) - \nabla F(U,V)\|_F \le 4\beta_0^2 \|(X,Y)-(U,V)\|_F. \tag{69}$$

By the triangular inequality, we have

$$\begin{aligned}&\|\nabla_X F(X,Y) - \nabla_X F(U,V)\|_F\\ &\quad\le \|\nabla_X F(X,Y) - \nabla_X F(U,Y)\|_F\\ &\qquad + \|\nabla_X F(U,Y) - \nabla_X F(U,V)\|_F.\end{aligned} \tag{70}$$

The first term of (70) can be bounded as follows

$$\begin{aligned}&\|\nabla_X F(X,Y) - \nabla_X F(U,Y)\|_F\\ &\quad= \|\mathcal{P}_\Omega(XY^T - M)Y - \mathcal{P}_\Omega(UY^T - M)Y\|_F\\ &\quad\le \|\mathcal{P}_\Omega(XY^T - M) - \mathcal{P}_\Omega(UY^T - M)\|_F\|Y\|_F\\ &\quad= \|\mathcal{P}_\Omega[(X-U)Y^T]\|_F\|Y\|_F\\ &\quad\le \|(X-U)Y^T\|_F\|Y\|_F\\ &\quad\le \|X-U\|_F\|Y\|_F^2\\ &\quad\le \|X-U\|_F\beta_0^2.\end{aligned}$$

The second term of (70) can be bounded as

$$\begin{aligned}&\|\nabla_X F(U,Y) - \nabla_X F(U,V)\|_F\\ &\quad= \|\mathcal{P}_\Omega(UY^T - M)Y - \mathcal{P}_\Omega(UV^T - M)V\|_F\\ &\quad\le \|\mathcal{P}_\Omega(M)(V-Y)\|_F + \|\mathcal{P}_\Omega(UY^T)Y - \mathcal{P}_\Omega(UV^T)V\|_F\\ &\quad\le \|\mathcal{P}_\Omega(M)(V-Y)\|_F + \|\mathcal{P}_\Omega(UY^T)Y - \mathcal{P}_\Omega(UY^T)V\|_F\\ &\qquad + \|\mathcal{P}_\Omega(UY^T)V - \mathcal{P}_\Omega(UV^T)V\|_F\\ &\quad\le \|\mathcal{P}_\Omega(M)\|_F\|V-Y\|_F + \|\mathcal{P}_\Omega(UY^T)\|_F\|Y-V\|_F\\ &\qquad + \|\mathcal{P}_\Omega[U(Y-V)^T]\|_F\|V\|_F\\ &\quad\le \|M\|_F\|V-Y\|_F + \|U\|_F\|Y\|_F\|Y-V\|_F\\ &\qquad + \|U\|_F\|Y-V\|_F\|V\|_F\\ &\quad\le 3\beta_0^2\|Y-V\|_F,\end{aligned}$$

where the last inequliaty follows from (68) and the fact that $\|M\|_F \le \sqrt{r}\Sigma_{\max} \overset{(15)}{=} \frac{1}{C_T\sqrt{r}}\beta_T^2 \le \beta_T^2 \le \beta_0^2$ (here the second last inequality follows from the fact that the numerical constant $C_T \ge 1$, and the last inequality follows from the assumption of Claim 2.1).

Plugging the above two bounds into (70), we obtain

$$\begin{aligned}&\|\nabla_X F(X,Y) - \nabla_X F(U,V)\|_F\\ &\quad\le \beta_0^2(\|X-U\|_F + 3\|Y-V\|_F).\end{aligned}$$

Similarly, we have

$$\begin{aligned}&\|\nabla_Y F(X,Y) - \nabla_Y F(U,V)\|_F\\ &\quad\le \beta_0^2(3\|X-U\|_F + \|Y-V\|_F).\end{aligned}$$

Combining the above two relations, we have (denote $\omega_1 \triangleq \|X-U\|_F, \omega_2 \triangleq \|Y-V\|_F$) the equation shown at the bottom of this page, which proves (69).

Next we prove

$$\|\nabla G(X,Y) - \nabla G(U,V)\|_F \le 54\rho\frac{\beta_0^2}{\beta_1^4}\|(X,Y)-(U,V)\|_F. \tag{71}$$

Denote

$$G_{1i}(X) \triangleq G_0\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right), \quad G_2(X) \triangleq G_0\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right), \tag{72}$$

then we have

$$\begin{aligned}\nabla G_{1i}(X) &= G_0'\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right)\frac{3\bar{X}^{(i)}}{\beta_1^2},\\ \nabla G_2(X) &= G_0'\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right)\frac{3X}{\beta_T^2},\end{aligned} \tag{73}$$

where $G_0'(z) = I_{[1,\infty]}(z)2(z-1)$ and $\bar{X}^{(i)}$ denotes a matrix with the i-th row being $X^{(i)}$ and the other rows being zero. Obviously $G_{1i}(X)$ is a matrix with all but the i-th row being zero. Recall that

$$G(X,Y) = \rho\sum_i G_{1i}(X) + \rho G_2(X) + f_0(Y),$$

where $f_0(Y)$ is a certain function of Y which we can ignore for now. Then we have

$$\begin{aligned}\nabla_X G(X,Y) &= \rho\sum_i \nabla G_{1i}(X) + \rho\nabla G_2(X)\\ &= \rho\sum_{i=1}^m G_0'\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right)\frac{3\bar{X}^{(i)}}{\beta_1^2} + \rho G_0'\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right)\frac{3X}{\beta_T^2},\end{aligned} \tag{74}$$

and, similarly,

$$\nabla_X G(U,V) = \rho\sum_i \nabla G_{1i}(U) + \rho G_2(U).$$

$$\begin{aligned}\|\nabla F(X,Y) - \nabla F(U,V)\|_F &= \sqrt{\|\nabla_X F(X,Y) - \nabla_X F(U,V)\|_F^2 + \|\nabla_Y F(X,Y) - \nabla_Y F(U,V)\|_F^2}\\ &\le \beta_0^2\sqrt{(\omega_1 + 3\omega_2)^2 + (3\omega_1+\omega_2)^2}\\ &\le 4\beta_0^2\sqrt{\omega_1^2+\omega_2^2}\\ &= 4\beta_0^2\|(X,Y)-(U,V)\|_F,\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\|\nabla_X G(X,Y) - \nabla_X G(U,V)\|_F \\
&\quad = \|\rho \sum_i [\nabla G_{1i}(X) - \nabla G_{1i}(U)] \\
&\qquad + \rho[\nabla G_2(X) - \nabla G_2(U)]\|_F \\
&\quad \le \|\rho \sum_i [\nabla G_{1i}(X) - \nabla G_{1i}(U)]\|_F \\
&\qquad + \rho\|\nabla G_2(X) - \nabla G_2(U)\|_F \\
&\quad = \rho \sqrt{\sum_i \|\nabla G_{1i}(X) - \nabla G_{1i}(U)\|_F^2} \\
&\qquad + \rho\|\nabla G_2(X) - \nabla G_2(U)\|_F, \qquad (75)
\end{aligned}$$

where the last equality is due to the fact that each $\nabla G_{1i}(X) - \nabla G_{1i}(U)$ is a matrix with all but the i-th row being zero. Denote

$$z_1 \triangleq \frac{3\|X\|_F^2}{2\beta_T^2}, \quad z_2 \triangleq \frac{3\|U\|_F^2}{2\beta_T^2}. \qquad (76)$$

Then by (76), (73) and the triangle inequality we have

$$\begin{aligned}
&\frac{\beta_T^2}{3}\|\nabla G_2(X) - \nabla G_2(U)\|_F \\
&\quad = \|G_0'(z_1)X - G_0'(z_2)U\|_F \\
&\quad \le |G_0'(z_1)|\|X - U\|_F + |G_0'(z_1) - G_0'(z_2)|\|U\|_F. \qquad (77)
\end{aligned}$$

By the definitions of z_1, z_2 in (76) and using $\|X\|_F \le \beta_0$, $\|Y\|_F \le \beta_0$, we have

$$\begin{aligned}
|z_1 - z_2| &= \frac{3}{2\beta_T^2}(\|X\|_F^2 - \|U\|_F^2) \\
&= \frac{3}{2\beta_T^2}(\|X\|_F + \|U\|_F)(\|X\|_F - \|U\|_F) \\
&\le \frac{3\beta_0}{\beta_T^2}\|X - U\|_F. \qquad (78)
\end{aligned}$$

According to (68) and the definitions of z_1, z_2 in (76), we have

$$\max\{z_1, z_2\} \le \frac{3}{2}\frac{\beta_0^2}{\beta_T^2}. \qquad (79)$$

We can bound the first and second order derivative of G_0 as follows:

$$G_0'(z) = I_{[1,\infty]}(z)2(z-1) \le 3\frac{\beta_0^2}{\beta_T^2}, \quad \forall z \in [0, \frac{3}{2}\frac{\beta_0^2}{\beta_T^2}], \qquad (80)$$

$$G_0''(z) = 2I_{[1,\infty]}(z) \le 2, \quad \forall z \in [0, \infty). \qquad (81)$$

By the mean value theorem and (81), we have

$$|G_0'(z_1) - G_0'(z_2)| \le 2|z_1 - z_2| \overset{(78)}{\le} \frac{6\beta_0}{\beta_T^2}\|X - U\|_F. \qquad (82)$$

Plugging (80) (with $z = z_1$) and (82) into (77), we obtain

$$\begin{aligned}
&\frac{\beta_T^2}{3}\|\nabla G_2(X) - \nabla G_2(U)\|_F \\
&\quad \le 3\frac{\beta_0^2}{\beta_T^2}\|X - U\|_F + \frac{6\beta_0}{\beta_T^2}\|X - U\|_F\|U\|_F \\
&\quad \le 9\frac{\beta_0^2}{\beta_T^2}\|X - U\|_F \\
&\quad \Longrightarrow \|\nabla G_2(X) - \nabla G_2(U)\|_F \le 27\frac{\beta_0^2}{\beta_T^4}\|X - U\|_F. \qquad (83)
\end{aligned}$$

Since $\|X^{(i)}\|_F \le \|X\|_F \le \beta_0$, $\|U^{(i)}\| \le \|U\|_F \le \beta_0$, by an argument analogous to that for (83), we can prove

$$\|\nabla G_{1i}(X) - \nabla G_{1i}(U)\|_F \le 27\frac{\beta_0^2}{\beta_1^4}\|X^{(i)} - U^{(i)}\|, \quad \forall\, i,$$

which further implies

$$\begin{aligned}
&\sqrt{\sum_i \|\nabla G_{1i}(X) - \nabla G_{1i}(U)\|^2} \\
&\quad \le 27\frac{\beta_0^2}{\beta_1^4}\sqrt{\sum_i \|X^{(i)} - U^{(i)}\|^2} = 27\frac{\beta_0^2}{\beta_1^4}\|X - U\|_F. \qquad (84)
\end{aligned}$$

Plugging (83) and (84) into (75), we obtain

$$\|\nabla_X G(X,Y) - \nabla_X G(U,V)\|_F \le 54\rho\frac{\beta_0^2}{\beta_1^4}\|X - U\|_F.$$

Similarly, we can prove

$$\begin{aligned}
\|\nabla_Y G(X,Y) - \nabla_Y G(U,V)\|_F &\le 54\rho\frac{\beta_0^2}{\beta_2^4}\|Y - V\|_F \\
&\le 54\rho\frac{\beta_0^2}{\beta_1^4}\|Y - V\|_F,
\end{aligned}$$

where the last inequality is due to $\beta_1 = \beta_T\sqrt{\frac{3\mu r}{m}} \le \beta_T\sqrt{\frac{3\mu r}{n}} = \beta_2$. Combining the above two relations yields (71).

Finally, we combine (69) and (71) to obtain

$$\begin{aligned}
&\|\nabla\tilde{F}(X,Y) - \nabla\tilde{F}(U,V)\|_F \\
&\le \|\nabla F(X,Y) - \nabla F(U,V)\|_F + \|\nabla G(X,Y) - \nabla G(U,V)\|_F \\
&\le \left(4\beta_0^2 + 54\rho\frac{\beta_0^2}{\beta_1^4}\right)\|(X,Y) - (U,V)\|_F,
\end{aligned}$$

which finishes the proof of Claim 2.1. □

Remark: If we further assume that the norm of each $X^{(i)}$ (resp. $Y^{(j)}$) is bounded by $O(\beta_1)$ (resp. $O(\beta_2)$), the Lipschitz constant can be improved to $4\beta_0^2 + 54\rho\frac{\beta_0^2}{\beta_T^4}$.

A.2 Solving the Subproblem of Algorithm 3

The subproblem of Algorithm 3 for the row vector $X^{(i)}$ is

$$\begin{aligned}
&\min_{X^{(i)}} \tilde{F}(X_k^{(1)}, \ldots, X_k^{(i-1)}, X^{(i)}, X_{k-1}^{(i+1)} \ldots, X_{k-1}^{(m)}, Y_{k-1}) \\
&\qquad + \frac{\lambda_0}{2}\|X^{(i)} - X_{k-1}^{(i)}\|^2.
\end{aligned}$$

For simplicity, denote $X^{(i)} = x_i, X_{k-1}^{(i)} = \bar{x}_i,\ X_k^{(j)} = x_j, 1 \le j \le i-1,\ X_{k-1}^{(j)} = x_j, i+1 \le j \le m,$ and $Y_{k-1}^{(j)} = y_j, 1 \le j \le n$. Then the above problem becomes

$$\min_{x_i} \tilde{F}(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m, y_1, \dots, y_n) + \frac{\lambda_0}{2}\|x_i - \bar{x}_i\|^2.$$

The optimal solution x_i^* to this subproblem satisfies the equation $\nabla_{x_i}\tilde{F} = 0$, i.e.

$$Ax_i - b + g(\|x_i\|)x_i = 0, \tag{85}$$

where $A = \sum_{j\in\Omega_i^x} y_j y_j^T + \lambda_0 I$ is a symmetric PD (positive definite) matrix, $b = \sum_{j\in\Omega_i^x} M_{ij} y_j + \lambda_0 \bar{x}_i$, and g is a function defined as

$$g(z) = \rho\frac{3}{\beta_1^2}G_0'(\frac{3z^2}{2\beta_1^2}) + \rho\frac{3}{\beta_T^2}G_0'(\frac{3(z^2+\xi_i)}{2\beta_T^2}),$$

in which $\xi_i = \sum_{j\ne i}\|x_j\|^2$ is a constant. Note that g has the following properties: a) $g(z) = 0$ when $z^2 \le \min\{\frac{2\beta_1^2}{3}, \frac{2\beta_T^2}{3} - \xi_i\}$; b) g is an increasing function in $[0, \infty)$. The equation (85) is equivalent to

$$x_i = (A + g(\|x_i\|)I)^{-1}b. \tag{86}$$

Suppose the eigendecomposition of A is $B\Lambda B^T$ and let $\Phi = B^T bb^T B$, then (86) implies

$$\begin{aligned}\|x_i\|^2 &= \|(A + g(\|x_i\|)I)^{-1}b\|^2 = \mathrm{Tr}((A + g(\|x_i\|)I)^{-2}bb^T)\\ &= \mathrm{Tr}((\Lambda + g(\|x_i\|)I)^{-2}\Phi) = \sum_{k=1}^r \frac{\Phi_{kk}}{(\Lambda_{kk} + g(\|x_i\|))^2},\\ &\Longrightarrow 1 = \frac{1}{\|x_i\|^2}\sum_{k=1}^r \frac{\Phi_{kk}}{(\Lambda_{kk} + g(\|x_i\|))^2},\end{aligned} \tag{87}$$

where Z_{kk} denotes the (k,k)-th entry of matrix Z. Since A and Φ are PSD (positive semidefinite) matrices, we have $\Phi_{kk} \ge 0, \Lambda_{kk} \ge 0$. The righthand side of (87) is a decreasing function of $\|x_i\|$, thus the equation (87) can be solved via a simple bisection procedure. After obtaining the norm of the optimal solution $z^* = \|x_i^*\|$, the optimal solution x_i^* can be obtained by (86), i.e.

$$x_i^* = (A + g(z^*)I)^{-1}b. \tag{88}$$

Similarly, the subproblem for $Y^{(j)}$ can also be solved by a bisection procedure.

B. Proof of Proposition 4.1

B.1 Matrix Norm Inequalities

We first prove some basic inequalities related to the matrix norms. These simple results will be used in the proof of Propositions 4.1 and 4.2.

Proposition B.1: If $A, B \in \mathbb{R}^{n_1\times n_2}$, then

$$\|A - B\|_2 \ge \sigma_{\min}(A) - \sigma_{\min}(B). \tag{89}$$

Proof: $\sigma_{\min}(A) = \min_{\|v\|=1}\|Av\| \le \min_{\|v\|=1}(\|Bv\| + \|(A-B)v\|) \le \min_{\|v\|=1}\|Bv\| + \|A-B\| = \sigma_{\min}(B) + \|A-B\|$.

Proposition B.2: For any $A \in \mathbb{R}^{n_1\times n_2}, B \in \mathbb{R}^{n_2\times n_3}$, we have

$$\sigma_{\min}(AB) \le \sigma_{\min}(A)\|B\|_2. \tag{90}$$

Proof: $\sigma_{\min}(AB) = \min_{v\in\mathbb{R}^{n_1\times 1}, \|v\|=1}\|v^T AB\| \le \min_{v\in\mathbb{R}^{n_1\times 1}, \|v\|=1}\|v^T A\|\|B\|_2 = \sigma_{\min}(A)\|B\|_2$.

Proposition B.3: Suppose $A, B \in \mathbb{R}^{n_1\times n_2}$ and $c_i A^{(i)} = B^{(i)}$, where $c_i \in \mathbb{R}$ and $|c_i| \le 1$, for $i = 1, \dots, n_1$ (recall that $Z^{(i)}$ denotes the i-th row of Z). Then

$$\|B\|_2 \le \|A\|_2.$$

Proof: For simplicity, denote $a_i \triangleq (A^{(i)})^T, b_i \triangleq (B^{(i)})^T$. Then

$$\begin{aligned}\|B\|_2^2 &= \max_{\|v\|=1}\|Bv\|^2 = \max_{\|v\|=1}\sum_i (b_i^T v)^2\\ &= \max_{\|v\|=1}\sum_i c_i^2(a_i^T v)^2 \le \max_{\|v\|=1}\sum_i (a_i^T v)^2 = \|A\|_2^2.\end{aligned}$$

Corollary B.1: Suppose $B \in \mathbb{R}^{n_1\times n_2}$ is a submatrix of $A \in \mathbb{R}^{m_1\times m_2}$, then

$$\|B\|_2 \le \|A\|_2. \tag{91}$$

Proof: By Proposition B.3, we have

$$\|(X_1, X_2)\|_2 \ge \|(X_1, 0)\|_2 = \|X_1\|_2.$$

Without loss of generality, suppose $A = \begin{bmatrix} B & B_1 \\ B_2 & B_3\end{bmatrix}$. Applying the above inequality twice, we get

$$\|A\|_2 \ge \|(B, B_1)\|_2 \ge \|B\|_2.$$

Proposition B.4: For any $A \in \mathbb{R}^{n_1\times n_2}, B \in \mathbb{R}^{n_2\times n_3}$, we have

$$\|AB\|_F \le \|A\|_2\|B\|_F, \tag{92a}$$
$$\|AB\|_2 \le \|A\|_2\|B\|_2. \tag{92b}$$

Further, if $n_1 \ge n_2$, then

$$\sigma_{\min}(A)\|B\|_F \le \|AB\|_F, \tag{93a}$$
$$\sigma_{\min}(A)\|B\|_2 \le \|AB\|_2. \tag{93b}$$

Proof: Assume the SVD of A is $A_1 D A_2$, where $A_1 \in \mathbb{R}^{n_1\times n_1}, A_2 \in \mathbb{R}^{n_2\times n_2}$ are orthonormal matrices and $D \in \mathbb{R}^{n_1\times n_2}$ has nonzero entries $D_{ii}, i = 1, \dots, \min\{n_1, n_2\}$. Note that

$$\sigma_{\min}(A) \le D_{ii} \le \|A\|_2, \quad \forall i.$$

Let $B' = A_2 B$ and suppose the i-th row of B' is b_i, $i = 1, \dots, n_2$, then

$$\|AB\|_F^2 = \|DA_2B\|_F^2 = \|DB'\|_F^2 = \sum_{i=1}^{\min\{n_1,n_2\}} D_{ii}^2\|b_i\|^2. \tag{94}$$

The the RHS (right hand side) can be bounded from above as

$$\begin{aligned}\sum_{i=1}^{\min\{n_1,n_2\}} D_{ii}^2\|b_i\|^2 &\le \|A\|_2^2 \sum_{i=1}^{\min\{n_1,n_2\}}\|b_i\|^2\\ &\le \|A\|_2^2\sum_{i=1}^{n_2} b_i^2 = \|A\|_2^2\|B'\|_F^2 = \|A\|_2^2\|B\|_F^2.\end{aligned}$$

Combining the above relation and (94) leads to (92a).

If $n_1 \geq n_2$, then $\min\{n_1, n_2\} = n_2$, and the RHS of (94) can be bounded from below as

$$\sum_{i=1}^{\min\{n_1,n_2\}} D_{ii}^2 \|b_i\|^2 = \sum_{i=1}^{n_2} D_{ii}^2 \|b_i\|^2 \geq \sigma_{\min}(A)^2 \sum_{i=1}^{n_2} \|b_i\|^2 = \sigma_{\min}(A)^2 \|B'\|_F^2 = \sigma_{\min}(A)^2 \|B\|_F^2.$$

Combining the above relation and (94) leads to (93a).

Next we prove the inequalities related to the spectral norm. We have

$$\|AB\|_2 = \|DA_2B\|_2 = \|DB'\|_2 = \max_{\|v\|\leq 1, v\in\mathbb{R}^{n_1\times 1}} \|v^T DB'\|. \tag{95}$$

Note that $\{v^T D \mid \|v\| \leq 1, v \in \mathbb{R}^{n_1\times 1}\} \subseteq \{u^T \mid u \in \mathbb{R}^{n_2\times 1}, \|u\| \leq \|A\|_2\}$, thus the RHS of (95) can be bounded from above as

$$\max_{\|v\|\leq 1, v\in\mathbb{R}^{n_1\times 1}} \|v^T DB'\| \leq \max_{u\in\mathbb{R}^{n_2\times 1}, \|u\|\leq\|A\|_2} \|u^T B'\| = \|A\|_2\|B'\|_2 = \|A\|_2\|B\|_2.$$

Combining the above relation and (95) leads to (92b).

If $n_1 \geq n_2$, then $\{u^T \mid u \in \mathbb{R}^{n_2\times 1}, \|u\| \leq \sigma_{\min}(A)\} \subseteq \{v^T D \mid \|v\| \leq 1, v \in \mathbb{R}^{n_1\times 1}\}$ (in fact, for any $\|u\| \leq \sigma_{\min}(A)$, let $v_i = u_i/D_{ii}, i = 1, \ldots, n_2$ and $v_i = 0, n_2 < i \leq n_1$, where v_i denotes the i-th entry of v, then $v^T D = u^T$ and $\|v\| \leq 1$). Thus the RHS of (95) can be bounded from below as

$$\max_{\|v\|\leq 1, v\in\mathbb{R}^{n_1\times 1}} \|v^T DB'\| \geq \max_{u\in\mathbb{R}^{n_2\times 1}, \|u\|\leq\sigma_{\min}(A)} \|u^T B'\| = \sigma_{\min}(A)\|B'\|_2 = \sigma_{\min}(A)\|B\|_2.$$

Combining the above relation and (95) leads to (93b). □

B.2 Proof of Proposition 4.1

Let M, X, Y satisfy the condition (47). First, we specify the choice of U, V. Suppose the SVD of M is $M = \hat{U}\Sigma\hat{V} = Q_1\tilde{\Sigma}Q_2^T$, where $Q_1 \in \mathcal{R}^{m\times m}, Q_2 \in \mathcal{R}^{n\times n}$ are unitary matrices, and $\tilde{\Sigma} = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$. Suppose $Q_1 = (Q_{11}, Q_{12}), Q_2 = (Q_{21}, Q_{22})$, where $Q_{11} = \hat{U} \in \mathcal{R}^{m\times r}, Q_{21} = \hat{V} \in \mathcal{R}^{n\times r}$ are incoherent matrices, and $Q_{12} \in \mathbb{R}^{m\times(m-r)}, Q_{22} \in \mathbb{R}^{n\times(n-r)}$. Let us write X, Y as

$$X = Q_1 \begin{pmatrix} X_1' \\ X_2' \end{pmatrix}, \quad Y = Q_2 \begin{pmatrix} Y_1' \\ Y_2' \end{pmatrix}, \tag{96}$$

where $X_1', Y_1' \in \mathbb{R}^{r\times r}, X_2' \in \mathbb{R}^{(m-r)\times r}, Y_2' \in \mathbb{R}^{(n-r)\times r}$. Define

$$U \triangleq Q_1 \begin{pmatrix} U_1' \\ 0 \end{pmatrix}, \quad V \triangleq Q_2 \begin{pmatrix} V_1' \\ 0 \end{pmatrix}, \tag{97}$$

where

$$U_1' = (1-\bar{\eta})X_1', \quad V_1' = \frac{1}{1-\bar{\eta}}\Sigma(X_1')^{-T},$$

in which

$$\bar{\eta} \triangleq \frac{d}{\Sigma_{\min}} \leq \frac{1}{11}.$$

The definition of V_1' is valid since X_1' is invertible (otherwise, $\operatorname{rank}(X_1'(Y_1')^T) \leq \operatorname{rank}(X_1') \leq r-1$, thus $d \geq \|\Sigma - X_1'(Y_1')^T\|_F \overset{(89)}{\geq} \Sigma_{\min} - \sigma_{\min}(X_1'(Y_1')^T) = \Sigma_{\min}$, which contradicts (47a). By this definition, we have

$$U_1'(V_1')^T = (1-\bar{\eta})X_1'(V_1')^T = \Sigma. \tag{98}$$

Now, we prove that U, V defined in (97) satisfy the requirement (48). The requirement (48a) $UV^T = M$ follows from (98) and (97). The requirement (48b) $\|U\|_F \leq (1-\frac{d}{\Sigma_{\min}})\|X\|_F$ can be proved as follows:

$$\|U\|_F = \|U_1'\|_F = (1-\frac{d}{\Sigma_{\min}})\|X_1'\|_F \leq (1-\frac{d}{\Sigma_{\min}})\|X\|_F.$$

As a side remark, the following variant of the requirement (48b) also holds:

$$\|U\|_2 \leq (1-\frac{d}{\Sigma_{\min}})\|X\|_2. \tag{99}$$

In fact, $\|U\|_2 = \|U_1'\|_2 = (1-\frac{d}{\Sigma_{\min}})\|X_1'\|_2 \overset{(91)}{\leq} (1-\frac{d}{\Sigma_{\min}})\left\|\begin{pmatrix} X_1' \\ X_2' \end{pmatrix}\right\|_2 = (1-\frac{d}{\Sigma_{\min}})\|X\|_2$.

To prove the requirement (48c), we first provide the bounds on $\|X_2'\|_F, \|V_1' - Y_1'\|_F, \|Y_2'\|_F$. Note that

$$\begin{aligned} d^2 &= \|M - XY^T\|_F^2 \\ &= \left\|\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} - Q_1^T XY^T Q_2\right\|_F^2 \\ &\overset{(96)}{=} \left\|\begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} X_1'(Y_1')^T & X_1'(Y_2')^T \\ X_2'(Y_1')^T & X_2'(Y_2')^T \end{pmatrix}\right\|_F^2 \\ &= \|\Sigma - X_1'(Y_1')^T\|_F^2 + \|X_1'(Y_2')^T\|_F^2 \\ &\quad + \|X_2'(Y_1')^T\|_F^2 + \|X_2'(Y_2')^T\|_F^2. \\ &\overset{(98)}{=} \|X_1'((1-\bar{\eta})V_1' - Y_1')^T\|_F^2 + \|X_1'(Y_2')^T\|_F^2 \\ &\quad + \|X_2'(Y_1')^T\|_F^2 + \|X_2'(Y_2')^T\|_F^2. \end{aligned} \tag{100}$$

Intuitively, since $\|X_1'\|_F, \|Y_1'\|_F$ are $O(1)$, we can upper bound $\|(1-\bar{\eta})V_1' - Y_1'\|_F, \|Y_2'\|_F, \|X_2'\|_F$ as $O(d)$. More rigorously, it follows from (100) that $d \geq \|X_1'((1-\bar{\eta})V_1' - Y_1')^T\|_F \overset{(93a)}{\geq} \sigma_{\min}(X_1')\|(1-\bar{\eta})V_1' - Y_1'\|_F$ and, similarly, $d \geq \sigma_{\min}(X_1')\|(Y_2')^T\|_F$, $d \geq \sigma_{\min}(Y_1')\|(X_2')^T\|_F$. These three inequalities imply

$$\begin{aligned} &\|(1-\bar{\eta})V_1' - Y_1'\|_F \leq \frac{d}{\sigma_{\min}(X_1')}, \\ &\|Y_2'\|_F \leq \frac{d}{\sigma_{\min}(X_1')}, \quad \|X_2'\|_F \leq \frac{d}{\sigma_{\min}(Y_1')}. \end{aligned} \tag{101}$$

We can lower bound $\sigma_{\min}(X_1')$ and $\sigma_{\min}(Y_1')$ as

$$\sigma_{\min}(X_1') \geq \frac{10\Sigma_{\min}}{11\beta_T}, \quad \sigma_{\min}(Y_1') \geq \frac{10\Sigma_{\min}}{11\beta_T}. \tag{102}$$

To prove (102), notice that (100) implies that $d \geq \|\Sigma - X_1'(Y_1')^T\|_F \geq \|\Sigma - X_1'(Y_1')^T\|_2 \overset{(89)}{\geq} \Sigma_{\min} - \sigma_{\min}(X_1'(Y_1')^T)$, which further implies

$$\sigma_{\min}(X_1'(Y_1')^T) \geq \Sigma_{\min} - d \geq \frac{10}{11}\Sigma_{\min}.$$

According to Proposition B.2, we have $\sigma_{\min}(X_1'(Y_1')^T) \leq \sigma_{\min}(X_1')\|Y_1'\|_2$. Combining this inequality with the above relation, we get $\sigma_{\min}(X_1')\|Y_1'\|_2 \geq \sigma_{\min}(X_1'(Y_1')^T) \geq 5\Sigma_{\min}/6$, which further implies

$$\sigma_{\min}(X_1') \geq \frac{10\Sigma_{\min}}{11\|Y_1'\|_2}. \tag{103}$$

Similarly, we have

$$\sigma_{\min}(Y_1') \geq \frac{10\Sigma_{\min}}{11\|X_1'\|_2}. \tag{104}$$

Plugging $\|Y_1'\|_2 \leq \|Y_1'\|_F \leq \|Y\|_F \leq \beta_T$ and similarly $\|X_1'\|_2 \leq \beta_T$ into (103) and (104), we obtain (102).

Combining (102) and (101), we obtain

$$\begin{aligned}&\max\{\|(1-\bar{\eta})V_1' - Y_1'\|_F, \|X_2'\|_F, \|Y_2'\|_F\}\\ &\quad \leq \frac{11}{10}\frac{d}{\Sigma_{\min}}\beta_T \leq \frac{1}{10}\beta_T. \end{aligned}\tag{105}$$

We can bound the norm of V_1' as

$$\begin{aligned}\|V_1'\|_F &= \frac{1}{1-\bar{\eta}}\|(1-\bar{\eta})V_1'\|_F\\ &\leq \frac{1}{1-\bar{\eta}}(\|(1-\bar{\eta})V_1' - Y_1'\|_F + \|Y_1'\|_F)\\ &\overset{(105)}{\leq} \frac{11}{10}\left(\frac{1}{10}\beta_T + \beta_T\right) \leq \left(\frac{11}{10}\right)^2 \beta_T. \end{aligned}\tag{106}$$

Combining this relation with (105), we have

$$\begin{aligned}\|V_1' - Y_1'\|_F &\leq \|(1-\bar{\eta})V_1' - Y_1'\|_F + \bar{\eta}\|V_1'\|_F\\ &\leq \frac{11}{10}\frac{d}{\Sigma_{\min}}\beta_T + \bar{\eta}\left(\frac{11}{10}\right)^2\beta_T \leq \frac{7\beta_T}{3\Sigma_{\min}}d.\end{aligned}$$

From (105) and the above relation we obtain

$$\begin{aligned}\|U - X\|_F &= \|X_2'\|_F \leq \frac{11\beta_T}{10\Sigma_{\min}}d \leq \frac{6\beta_T}{5\Sigma_{\min}}d,\\ \|V - Y\|_F &= \sqrt{\|V_1' - Y_1'\|_F^2 + \|Y_2'\|_F^2}\\ &\leq \sqrt{\left(\frac{7}{3}\right)^2 + \left(\frac{11}{10}\right)^2}\frac{\beta_T}{\Sigma_{\min}}d \leq \frac{3\beta_T}{\Sigma_{\min}}d,\end{aligned}$$

which finishes the proof of the requirement (48c).

As a side remark, the requirement (48c) can be slightly improved to

$$\|U - X\|_F \leq \frac{6\|Y\|_2}{5\Sigma_{\min}}d, \quad \|V - Y\|_F \leq \frac{3\|X\|_2}{\Sigma_{\min}}d. \tag{107}$$

In fact, plugging $\|X_1'\|_2 \overset{(91)}{\leq} \|\begin{pmatrix}X_1'\\ X_2'\end{pmatrix}\|_2 = \|X\|_2$ and similarly $\|Y_1'\|_2 \leq \|Y\|_2$ into (103) and (104), we obtain $\sigma_{\min}(X_1') \geq \frac{5\Sigma_{\min}}{6\|Y\|_2}$, $\sigma_{\min}(Y_1') \geq \frac{5\Sigma_{\min}}{6\|X\|_2}$. Combining with (101), we obtain (107). This inequality will be used in the proof of Claim 5.2 in Appendix D.1.

At last, we prove the requirement (48d). By the definitions of U, V in (97), we have

$$\begin{aligned}U &= (Q_{11}, Q_{12})\begin{pmatrix}U_1'\\ 0\end{pmatrix} = Q_{11}U_1',\\ V &= (Q_{21}, Q_{22})\begin{pmatrix}V_1'\\ 0\end{pmatrix} = Q_{21}V_1'.\end{aligned}\tag{108}$$

The assumption that M is μ-incoherent implies

$$\|Q_{11}^{(i)}\|^2 = \|\hat{U}^{(i)}\|^2 \leq \frac{r\mu}{m}, \quad \|Q_{21}^{(i)}\|^2 = \|\hat{V}^{(j)}\|^2 \leq \frac{r\mu}{n}, \quad \forall i, j.$$

Notice the following fact: for any matrix $A \in \mathbb{R}^{K\times r}$, $B \in \mathbb{R}^{r\times r}$, where $K \in \{m, n\}$, we have

$$\|(AB)^{(i)}\|^2 = \|A^{(i)}B\|^2 \leq \|A^{(i)}\|^2\|B\|_F^2.$$

Therefore, we have (using the fact $\|U_1'\|_F \leq \|X_1'\|_F \leq \|X\|_F \leq \beta_T$ and (106))

$$\begin{aligned}\|U^{(i)}\|^2 &= \|(Q_{11}U_1')^{(i)}\|^2 \leq \|Q_{11}^{(i)}\|^2\|U_1'\|_F^2 \leq \frac{r\mu}{m}\beta_T^2;\\ \|V^{(j)}\|^2 &= \|(Q_{21}V_1')^{(j)}\|^2 \leq \frac{r\mu}{n}\|V_1'\|_F^2\\ &\overset{(106)}{\leq} \left(\frac{11}{10}\right)^4\frac{r\mu}{n}\beta_T^2 \leq \frac{3}{2}\frac{r\mu}{n}\beta_T^2,\end{aligned}\tag{109}$$

which finishes the proof the requirement (48d).

We will first reduce Proposition 4.2 to Proposition C.1 for $r \times r$ matrices in Section V-C. This reduction is rather trivial, and the major difficulty lies in Proposition C.1. For general r, the proof of Proposition C.1 is rather involved. We will give the overview of the main proof ideas in Appendix C.2. Most readers can skip Appendix C.1.

C. Proof of Proposition 4.2

C.1 Transformation to a Simpler Problem

We first transform the problem to a simpler problem that only involves $r\times r$ matrices. In particular, we will show that to prove Proposition 4.2 we only need to prove Proposition C.1.

Similar to the proof of Proposition 4.1, we use $Q_1 \in \mathbb{R}^{m\times m}$, $Q_2 \in \mathbb{R}^{n\times n}$ to denote the SVD factors of M (Q_1 and Q_2 are unitary matrices), and write X, Y as

$$X = Q_1\begin{pmatrix}X_1'\\ X_2'\end{pmatrix}, \quad Y = Q_2\begin{pmatrix}Y_1'\\ Y_2'\end{pmatrix}.$$

Define

$$U = Q_1\begin{pmatrix}U_1'\\ 0\end{pmatrix}, \quad V = Q_2\begin{pmatrix}V_1'\\ 0\end{pmatrix}, \tag{110}$$

where $U_1' \in \mathbb{R}^{r\times r}$ and $V_1' \in \mathbb{R}^{r\times r}$ are to be determined.

We can convert the conditions on U, V to the conditions on U_1', V_1'. As proved in Appendix A.2 (combining (101) and (102)),

$$\|X_2'\|_F \leq \frac{6\beta_T}{5\Sigma_{\min}}d, \quad \|Y_2'\|_F \leq \frac{6\beta_T}{5\Sigma_{\min}}d. \tag{111}$$

Obviously, the condition (49a) implies the following condition on X_1', Y_1':

$$d' \triangleq \|\Sigma - (X_1')(Y_1')^T\| \leq \frac{\Sigma_{\min}}{C_d r}. \tag{112}$$

Using (111) and the facts $\|X\|_F = \sqrt{\|X_1'\|_F^2 + \|X_2'\|_F^2}$ and $\|Y\|_F = \sqrt{\|Y_1'\|_F^2 + \|Y_2'\|_F^2}$, the condition (49b) implies the following condition on X_1', Y_1':

$$\sqrt{\frac{3}{5}}\beta_T \leq \|X_1'\|_F \leq \beta_T, \quad \sqrt{\frac{3}{5}}\beta_T \leq \|Y_1'\|_F \leq \beta_T. \tag{113}$$

We have the following proposition.

Proposition C.1: There exist numerical constants C_d, C_T such that: if $X_1', Y_1' \in \mathbb{R}^{r\times r}$ satisfy (112) *and* (113)*, where $\beta_T = \sqrt{C_T r \Sigma_{\max}}$, then there exist $U_1' \in \mathbb{R}^{r\times r}, V_1' \in \mathbb{R}^{r\times r}$ such that*

$$U_1'(V_1')^T = \Sigma, \tag{114a}$$

$$\|U_1'\|_F \le \|X_1'\|_F,$$
$$\|V_1'\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|Y_1'\|_F, \tag{114b}$$

$$\|U_1' - X_1'\|_F \|V_1' - Y_1'\|_F \le 63\sqrt{r}\frac{\beta_T^2}{\Sigma_{\min}^2}d^2,$$
$$\max\{\|U_1' - X_1'\|_F, \|V_1' - Y_1'\|_F\} \le \frac{58}{7}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d. \tag{114c}$$

We claim that Proposition C.1 implies Proposition 4.2. Since we have already proved that the conditions of Proposition 4.2 imply the conditions of Proposition C.1, we only need to prove that the conclusion of Proposition C.1 implies the conclusion of Proposition 4.2. In other words, we only need to show that if U_1', V_1' satisfy (114), then they satisfy the requirements (50).

The requirement (50a) $UV^T = M$ follows directly from (114a) and the definition of U, V in (110). The requirement (50b) can be proved as $\|V\|_F = \|V_1'\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|Y_1'\|_F \le (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F$ and $\|U\|_F = \|U_1'\|_F \le \|X\|_F$. Analogous to (109), the requirement (50d) can be proved as $\|V^{(j)}\|^2 = \|(Q_{21}V_1')^{(j)}\|^2 \le \frac{r\mu}{n}\|V_1'\|_F^2 \le \frac{r\mu}{n}\beta_T^2$ and, similarly, $\|U^{(i)}\|^2 \le \frac{r\mu}{m}\beta_T^2$. At last, we prove the requirement (50c). The first relation in (50c) can be proved as

$$\begin{aligned}
&\|U - X\|_F\|V - Y\|_F \\
&= \sqrt{\|U_1' - X_1'\|_F^2 + \|X_2'\|_F^2}\sqrt{\|V_1' - Y_1'\|_F^2 + \|Y_2'\|_F^2} \\
&= (\| U_1' - X_1'\|_F^2\|V_1' - Y_1'\|_F^2 + \|X_2'\|_F^2\|V_1' - Y_1'\|_F^2 \\
&\qquad + \|U_1' - X_1'\|_F^2\|Y_2'\|_F^2 + \|X_2'\|_F^2\|Y_2'\|_F^2\Big)^{1/2} \\
&\overset{(111),(114c)}{\le} \sqrt{r}\frac{\beta_T^2}{\Sigma_{\min}^2}d^2\sqrt{63^2 + (\frac{6}{5})^2(\frac{58}{7})^2 + (\frac{58}{7})^2(\frac{6}{5})^2 + (\frac{6}{5})^4}, \\
&< 65\sqrt{r}\frac{\beta_T^2}{\Sigma_{\min}^2}d^2,
\end{aligned}$$

where in the second last inequality we also use the fact $d' \le d$. The second relation in (50c) can be proved by

$$\begin{aligned}
\|U - X\|_F &= \sqrt{\|U_1' - X_1'\|_F^2 + \|X_2'\|_F^2} \\
&\overset{(111),(114c)}{\le} \sqrt{(\frac{6}{5})^2 + (\frac{58}{7})^2}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d \le \frac{17}{2}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d
\end{aligned}$$

and a similar inequality for $\|V - Y\|_F$.

C.2 Preliminary Analysis for the Proof of Proposition C.1

We first give a more intuitive explanation of what we want to prove, by relating the result to "preconditioning". Then we analyze two simple examples for $r = 2$ to get some ideas on how to approach the problem. Next we discuss how to extend the ideas to general r. To simplify the notations, from now on, we use X, Y, U, V, d to replace $X_1', Y_1', U_1', V_1', d'$ in Proposition (C.1).

C.2.1 Perturbation Analysis for Preconditioning

We claim that Proposition C.1 is closely related to "preconditioning", which refers to reducing the condition number (by preprocessing) in numerical linear algebra.

Proposition C.2 (Informal): Suppose $X \in \mathbb{R}^{r\times r}$ is non-singular and $\|X\|_F = \|X^{-1}\|_F \ge C\sqrt{r}$ where $C \ge 10$ is a constant. For any $d' \le \mathcal{O}(1/r^{1.5})$, there exists $U \in \mathbb{R}^{r\times r}$ such that $\|U\|_F = \|X\|_F$, $\|U^{-1}\|_F \le (1 - d')\|X^{-1}\|_F$ and $\max\{\|U - X\|_F, \|U^{-1} - X^{-1}\|_F\} \le \mathcal{O}(d'r^{1.5})$.

We will argue later that Proposition C.2 is a simple version of Proposition C.1.

We explain why this proposition can be understood as perturbation analysis for perconditioning. Assume X has singular values $\sigma_1 \ge \cdots \ge \sigma_r > 0$, then $\|X\|_F^2 = \sum_i \sigma_i^2$ and $\|X^{-1}\|_F^2 = \sum_i \frac{1}{\sigma_i^2}$. By Cauchy-Schwartz inequality $\|X\|_F^2\|X^{-1}\|_F^2 \ge r^2$, and the equality holds iff $\sigma_1 = \cdots = \sigma_r$, i.e., X has a condition number 1. In other words, if $\|X\|_F = \|X^{-1}\|_F = \sqrt{r}$, then X has the minimal condition number 1. In the assumption $\|X\|_F = \|X^{-1}\|_F \ge C\sqrt{r}$, C can be viewed as a measure of the ill-conditioned-ness of X (different from the condition number σ_1/σ_r but related). Prop. C.2 simply says that we can perturb X to make X better-conditioned.

Prop. C.2 itself is not difficult to prove. In fact, without loss of generality we can assume X is a diagonal matrix (by left and right multiplying X by its singular vector matrices). Then the problem reduces to the following problem: assume $\sum_i \sigma_i^2 = \sum_i \frac{1}{\sigma_i^2} \ge C^2 r$, perturb σ_i's so that the $\sum_i \sigma_i^2$ does not change while $\sum_i \frac{1}{\sigma_i^2}$ increases. This is a rather easy problem. Nevertheless, for the original desired result Prop. C.1 we cannot assume X is diagonal. In Section V-C we will analyze the problem without assuming X is diagonal.

To show the connection of Prop. C.2 and Prop. C.1, we first simplify the statement of Prop. C.1.

Proposition C.3 (Simpler Version of Proposition C.1): Suppose $X, Y, \Sigma \in \mathbb{R}^{r\times r}$ are non-singular and Σ is diagonal. If $\|XY^T - \Sigma\|_F = d \le \mathcal{O}(\Sigma_{\min}/r)$ and $\|X\|_F = \|Y\|_F = \beta \ge C\sqrt{r\Sigma_{\max}}$, then we can find a factorization $\Sigma = UV^T$ such that $\max\{\|U - X\|_F, \|V - Y\|_F\} \le \mathcal{O}(\sqrt{r}d\beta/\Sigma_{\min})$ and $\|U\|_F \le \|X\|_F$, $\|V\|_F \le \|Y\|_F$.

There are a few differences with Prop. C.1: i) In Prop. C.1 we assume $\|X\|_F, \|Y\|_F \in [\sqrt{0.6}\beta_T, \beta_T]$, but by simply scaling X, U, Y, V we can assume $\|X\|_F = \|Y\|_F$ as in the above proposition; ii) here we only require $\|V\|_F \le \|Y\|_F$, instead of $\|V\|_F \le (1 - d/\Sigma_{\min})\|Y\|_F$ in (114b); iii) in Prop. C.1 there is an extra bound of $\|U - X\|_F\|V - Y\|_F$. Nevertheless, these differences are not essential and do not affect the proof too much.

Now let us consider a special case and show how to reduce Prop. C.3 to Prop. C.2. This part is mainly for the purpose of rigorous derivation and we suggest first-time readers jump to Section V-C. The special case we consider is $\Sigma = I$ and $XY^T = (1 - d/\sqrt{r})I$, where $d \le \mathcal{O}(1/r)$. Let $d' = d/\sqrt{r} \le \mathcal{O}(1/r^{1.5})$, then $Y^T = (1 - d/\sqrt{r})X^{-1} = (1 - d')X^{-1}$. The condition of Prop. C.3 becomes

$$\|X\|_F = \|X^{-1}\|_F(1 - d') = \beta. \tag{115}$$

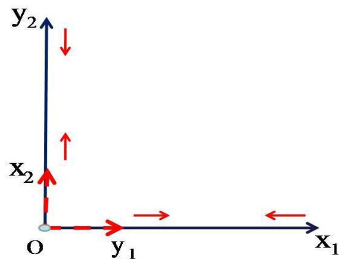

Fig. 3. Illustration of the first example. $X = (x_1^T, x_2^T) = \text{Diag}\,(x_{11}, x_{22})$, $Y = (y_1^T, y_2^T) = \text{Diag}\,(y_{11}, y_{22})$, where $x_{11} = y_{22} \gg x_{22} = y_{11}$ and $x_{11}y_{11} = x_{22}y_{22} = 1 - d/\sqrt{2}$. We use the following operation to define U, V: shrink x_1 and extend x_2 to obtain U, while keeping the norm invariant (i.e. $\|U\|_F = \|X\|_F$); shrink y_2 and extend y_1 to obtain V, while keeping the norm invariant (i.e. $\|V\|_F = \|Y\|_F$). We can prove that there exists an operation such that $u_{ii}v_{ii} = 1 > x_{ii}y_{ii}, i = 1, 2$.

One requirement of Prop. C.3 becomes $\|U\|_F \le \|X\|_F, \|U^{-1}\|_F \le \|Y\|_F = \|X^{-1}\|_F(1 - d')$. The distance bound in Prop. C.3 is $\mathcal{O}(\sqrt{r}d\beta/\Sigma_{\min})$, which becomes $\mathcal{O}(d'r^{1.5})$ under the new parameter setting. By a similar scaling technique, i.e. scaling X, U by $1/\sqrt{1-d'}$ and Y, V by $\sqrt{1-d'}$, we can replace the condition (115) by

$$\|X\|_F = \|Y\|_F = \beta\sqrt{\frac{1}{1-d'}} \ge C\sqrt{r}.$$

Note that rigorously speaking the bound should be $C\sqrt{r}/\sqrt{1-d'}$, but since $1/(1-d') \le 1/(1-1/r^{1.5}) \in [1/(1-1/2^{1.5}), 1]$, the contribution of $1/\sqrt{1-d'}$ is just a numerical constant which can be absorbed into C. Now the problem becomes: assume $\|X\|_F = \|X^{-1}\|_F \ge C\sqrt{r}$, find U such that $\|U\|_F \le \|X\|_F$, $\|U^{-1}\|_F \le \|X^{-1}\|_F(1-d')$ and $\max\{\|U - X\|_F, \|U^{-1} - X^{-1}\|_F\} \le \mathcal{O}(d'r^{1.5})$, where $d' \le \mathcal{O}(1/r^{1.5})$. By slightly strengthening the requirement $\|U\|_F \le \|X\|_F$ to $\|U\|_F = \|X\|_F$, we obtain Prop. C.2.

C.2.2 Two Motivating Examples

We denote the i-th row of X, Y as x_i, y_i, respectively. In the first example (see Figure 3), we set $r = 2$, $\Sigma = I$ (which implies $\Sigma_{\min} = \Sigma_{\max} = 1$), $d = 1/(C_d r)$ and

$$\begin{aligned} X &= \text{Diag}\,(x_{11}, x_{22}) = \text{Diag}\left(C, \frac{1-d/\sqrt{2}}{C}\right), \\ Y &= \text{Diag}\,(y_{11}, y_{22}) = \text{Diag}\left(\frac{1-d/\sqrt{2}}{C}, C\right), \end{aligned} \tag{116}$$

where $C > 1$ is to be determined, and $\text{Diag}(w_1, w_2)$ denotes a 2×2 diagonal matrix with diagonal entries w_1, w_2. In this setting $\beta_T = \sqrt{rC_T\Sigma_{\max}} = \sqrt{2C_T}$ is a large constant. Condition (112) holds since $\|XY^T - \Sigma\|_F = \|(1-d/\sqrt{2})I - I\|_F = d = 1/(C_d r)$. Note that $\|X\|_F = \|Y\|_F = \sqrt{C^2 + \frac{(1-d/\sqrt{2})^2}{C^2}} \approx C$, thus there exists $C \in [\sqrt{3/5}\beta_T, \beta_T]$ so that (113) holds.

How should we define $U = \text{Diag}(u_{11}, u_{22}), V = \text{Diag}(v_{11}, v_{22})$ so that (114) holds? Due to the "symmetry" of X and Y in this example (by symmetry we mean $x_{11} = y_{22}, x_{22} = y_{11}$), we choose U, V such that $u_{11} = v_{22}, u_{22} = v_{11}$. Then the requirements (114a) and (114b) reduce to:

$$\begin{aligned} u_{11}u_{22} = 1 &= \frac{x_{11}x_{22}}{1-d/\sqrt{2}}, \\ u_{11}^2 + u_{22}^2 &\le x_{11}^2 + x_{22}^2. \end{aligned} \tag{117}$$

It can be easily shown that there exist u_{11}, u_{22} satisfying (117). In fact, define $R = \|X\|_F = \sqrt{x_{11}^2 + x_{22}^2}$ and let a point (w_1, w_2) move along the circle $\{(w_1, w_2) \mid w_1^2 + w_2^2 = R^2\}$ from (x_{11}, x_{22}) to $(R/\sqrt{2}, R/\sqrt{2})$. During this process, the norm of (w_1, w_2) does not change and the product w_1w_2 monotonically increases from $x_{11}x_{22}$ to $R^2/2$. Therefore, there exist u_{11}, u_{22} satisfying (117) as long as $R^2/2 > x_{11}x_{22}/(1-d/\sqrt{2})$. This inequality is equivalent to $(1-d/\sqrt{2})(x_{11}^2 + x_{22}^2)/2 > x_{11}x_{22}$, which can be simplified to $(1-d/\sqrt{2})(x_{11}-x_{22})^2 > \sqrt{2}dx_{11}x_{22} = \sqrt{2}d(1-d/\sqrt{2})$, or equivalently, $(x_{11}-x_{22})^2 > \sqrt{2}d$. The last inequality holds when $x_{11} - x_{22} = C - (1-d/\sqrt{2})/C$ is large enough (i.e. C is large enough).

To summarize, we will increase the small entry x_{22} (resp. y_{11}) and decrease the large entry x_{11} (resp. y_{22}) to obtain a more balanced diagonal matrix U (resp. V), which has the same norm as X (resp. Y). The percentage of increase in the small entry x_{22} (resp. y_{11}) will be much larger than the percentage of decrease in the large entry x_{11} (resp. y_{22}), thus the products $x_{22}y_{22}$ and $x_{11}y_{11}$ will increase; in other words, the product UV^T of the more balanced matrices U, V will have larger entries than XY^T.

Note that the above idea of shrinking/extending works when there is a large imbalance in the lengths of the rows of X, Y, regardless of whether X, Y are diagonal matrices or not. By the assumption that $\|X\|_F$ and $\|Y\|_F$ are large, we know that there must be a row of X (resp. Y) that has large norm (here "large" means much larger than $1/\sqrt{r}$); however, it is possible that all rows of X and Y have large norm and there is no imbalance in terms of the lengths of the rows. See below for such an example.

In the second example (see Figure 4), we still set $r = 2$, $\Sigma = I$, $d = 1/(C_d r)$. Suppose $X = (x_1^T, x_2^T)$, $Y = (y_1^T, y_2^T)$. We define $x_1 = (C, 0), x_2 = (-C\sin\alpha, C\cos\alpha)$ and $y_1 = (C\cos\alpha, C\sin\alpha), y_2 = (0, C)$, where C is a large constant, and $\alpha \in (0, \pi/2)$ is chosen so that

$$C^2\cos\alpha = 1 - d/\sqrt{2}. \tag{118}$$

When C is large, $\alpha \approx \arccos(1/C^2)$ is also large (i.e. close to $\pi/2$). Condition (112) holds since $\|XY^T - \Sigma\|_F = \|C^2\cos\alpha I - I\|_F = \|(1-d/\sqrt{2})I - I\|_F = d = 1/(C_d r)$. Note that $\|X\|_F = \|Y\|_F = \sqrt{2}C$, so we can choose $C = \beta_T/\sqrt{2} = \sqrt{2C_T}/\sqrt{2} = \sqrt{C_T}$ so that (113) holds.

How should we choose $U = (u_1^T, u_2^T)$, $V = (v_1^T, v_2^T)$ so that (114) holds? The idea for the first example no longer works since it requires that the difference of $\|x_1\|$ and $\|x_2\|$ (resp. $\|y_1\|$ and $\|y_2\|$) is large; however, in this example, $\|x_1\| - \|x_2\| = \|y_1\| - \|y_2\| = 0$. The key idea for this example is to use rotation. Rotating a vector does not change the norm, so requirement (113) will not be violated if u_i (resp. v_i) is obtained by rotating x_i(resp. y_i). For simplicity, we rotate y_1, x_2 to obtain v_1, u_2 respectively and let

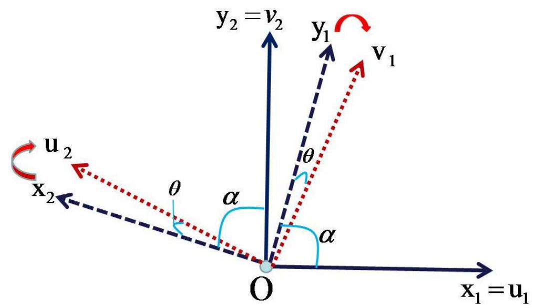

Fig. 4. Illustration of the second example. $X = (x_1^T, x_2^T)$, $Y = (y_1^T, y_2^T)$, where $x_1 = (C, 0), x_2 = (-C\sin\alpha, C\cos\alpha)$ and $y_1 = (C\cos\alpha, C\sin\alpha), y_2 = (0, C)$, where C is a large constant. Choose α so that $C^2\cos\alpha = 1 - d/\sqrt{2}$. We use the following operation to define $U = (u_1^T, u_2^T)$, $V = (v_1^T, v_2^T)$: rotate y_1 (resp. x_2) by angle θ to obtain v_1(resp. u_2), and let $u_2 = x_2, v_1 = y_1$. Here the angle of rotation θ is chosen so that $\langle u_1, v_1\rangle = \langle u_2, v_2\rangle = 1$.

$u_1 = x_1, v_2 = y_2$ (see Figure 4). Note that y_1 and x_2 should be rotated by the same angle as v_1 should be orthogonal to u_2 (since the off-diagonal entries of UV^T are zero). To increase the inner product $\langle x_i, y_i\rangle$ from $1 - d/\sqrt{2}$ to 1, we need to decrease the angle of x_i and y_i, thus y_1 (resp. x_2) should be rotated towards x_1(resp. y_2). Finally, let us specify the angle of rotation $\theta \triangleq \angle(y_1, v_1) = \angle(x_2, u_2)$. The requirement $\langle u_1, v_1\rangle = 1$ is equivalent to $1 = \|u_1\|\|v_1\|\cos\angle(u_1, v_1) = \|x_1\|\|y_1\|\cos(\alpha - \theta)$, which can be rewritten as

$$1 = C^2\cos(\alpha - \theta). \tag{119}$$

The right-hand side of (119) is an increasing function of θ, ranging from $C^2\cos(\alpha) \overset{(118)}{=} 1 - d/\sqrt{2}$ to C^2 for $\theta \in [0, \alpha]$. Since 1 lies in the range $[1 - d/\sqrt{2}, C^2]$, there exists a unique θ so that (119) holds. One can further verify the requirement (114c), i.e. the difference of X (resp. Y) and U (resp. V) is small. As a rough summary, we rotate x_i, y_i to obtain u_i, v_i when the angle of x_i and y_i is large. This operation does not change the norm and can increase the inner product $\langle x_i, y_i\rangle$ to the desired amount (1 in this case).

C.2.3 Proof Ideas of Proposition C.1

In the above two examples, we have used two different operations: one is based on shrinking/extending, and the other is based on rotation. As we mentioned before, the first operation cannot deal with the second example; also, it is obvious that the second operation cannot deal with the first example (the angle between x_i and y_i is zero, so rotation only decreases the inner product). Therefore, both operations are necessary.

Are these two operations sufficient? Fortunately, the answer is yes for the case that XY^T is diagonal and $\langle x_i, y_i\rangle \le \Sigma_i$ (we need extra effort to reduce the general problem to this case). When all the angles between x_i and y_i are smaller than a constant $\bar{\alpha}$, there must be some kind of imbalance in the lengths of x_i, y_i's (to illustrate this, if all $\|x_i\| = \|y_i\|$, then $\|x_i\|^2 = \|x_i\|\|y_i\| \approx \Sigma_i/\cos\angle(x_i, y_i) \le \Sigma_i/\cos(\bar{\alpha})$, which implies $\|X\|_F^2 \lesssim r\Sigma_{\max}/\cos(\bar{\alpha}) \ll \frac{3}{5}C_T r\Sigma_{\max} = \frac{3}{5}\beta_T^2$ for large enough C_T, a contradiction to (112)). Thus we can use the first operation (i.e. shrinking/extending the vectors x_i, y_i's) to obtain the desired U, V. When all the angles between x_i and y_i are larger than a constant $\bar{\alpha}$, we can use the second operation (i.e. rotating the vectors x_i, y_i's) to obtain the desired U, V. In general, some angles may be larger than $\bar{\alpha}$ and others may be smaller, then a natural solution is to use the two operations *simultaneously*: use the first operation for the pairs (x_i, y_i) with small angles and the second operation for those with large angles.

We had a proof using the two operations simultaneously, but the bounds on $\|U - X\|_F, \|V - Y\|_F$ have a large exponent of r. In the following subsection, we present a different proof that does not use the two operations simultaneously, but only use one of the two operations. The basic proof framework is summarized as follows. We first define $\hat{Y}$ so that $X\hat{Y} = \Sigma$; in other words, we try to satisfy the requirement (114a) first. Then we try to modify $\hat{Y}$ to satisfy the requirement (114b). In particular, we need to reduce the norm of $\hat{Y}$ and keep the norm of X unchanged, while maintaining the relation $X\hat{Y}^T = \Sigma$. We consider two cases: in Case 1, "most" angles between X and $\hat{Y}$ are smaller than $\bar{\alpha}$, and using the first operation (shrinking/extending) can obtain the desired U, V; in Case 2, "most" angles between X and $\hat{Y}$ are larger than $\bar{\alpha}$, and using the second operation (rotation) can obtain the desired U, V (see (127) for a precise definition of Case 1 and Case 2). The difference of this proof framework and the previous one is the following. In our previous proof framework, we need to take into account every pair x_i, y_i so that its inner product is modified to Σ_i, thus two operations have to be applied simultaneously. In contrast, in this new proof framework, $\langle x_i, \hat{y}_i\rangle$ is already Σ_i, and we only need to worry about the "overall" requirement that $\|\hat{Y}\|_F$ should be reduced, thus dealing only with the pairs with small angles (or only with the pairs with large angles) is enough to satisfy the requirement.

Finally, we would like to mention that when Σ is an identity matrix, the proof can be rather simple. In fact, in this case one can assume X to be diagonal by proper orthonormal transformation, and then assume Y to be diagonal since the off-diagonal entries are small. By just using the first operation (scaling of the diagonal entries), we can construct the desired U, V and the proof is similar to that in Appendix C.3.1. When Σ is not a diagonal matrix, we can replace X, Y by $XQ, Q^{-1}Y$ where Q is orthonormal, but that only simplifies X to a upper triangular matrix, a condition seems not very helpful. It seems that the second operation has to be used and the proof becomes more involved.

C.3 Proof of Proposition C.1

As mentioned earlier, to simplify the notations, we use X, Y, U, V, d to replace $X_1', Y_1', U_1', V_1', d'$ in Proposition (C.1). Throughout the proof, we choose

$$C_T = 20, \tag{120}$$

and $C_d = 108$, which implies

$$\frac{d}{\Sigma_{\min}} \le \frac{1}{108r}. \tag{121}$$

There are two "hard" requirements on U, V: (114a) and (114b). Our construction of U, V can be viewed as a two-step

TABLE VIII
OPERATION 1

Operation 1: Shrinking and Extending

Input: $x_k, \hat{y}_k, k = 1, \ldots, r$.

Output: $u_k, v_k, k = 1, \ldots, r$.

Procedure:

(i) For each $j \le s$, keep $x_j, \hat{y}_j$ unchanged, i.e.

$$u_j \triangleq x_j,\ v_j \triangleq \hat{y}_j,\ \ j = 1, \ldots, s. \tag{131}$$

(ii) For each $i \in \{s+1, \ldots, K\}$, extend x_i to obtain u_i and shrink $\hat{y}_i$ to obtain v_i. For each $i \ge K+1$, shrink x_i to obtain u_i and extend $\hat{y}_i$ to obtain v_i. More specifically,

$$u_i \triangleq \frac{x_i}{1-\epsilon_i}, v_i \triangleq \hat{y}_i(1-\epsilon_i),\ \text{where } \epsilon_i = \begin{cases} 7\bar{\eta} & i \le K, \\ -4.5\bar{\eta} & i \ge K+1, \end{cases} \quad i = s+1, s+2, \ldots, r, \tag{132}$$

in which

$$\bar{\eta} \triangleq \frac{d}{\Sigma_{\min}} \ge \eta. \tag{133}$$

approach, whereby we satisfy one requirement in each step. In Step 1, we construct

$$\hat{Y} = \Sigma(\Sigma + D)^{-T} Y, \quad \text{where } D \triangleq XY^T - \Sigma,$$

then

$$X\hat{Y}^T = (XY^T)(XY^T)^{-1}\Sigma = \Sigma,$$

i.e. the first requirement is satisfied. Since the new $\hat{Y}$ may have higher norm than $\|Y\|_F$, in Step 2 we modify $X, \hat{Y}$ to U, V so that the product does not change, and $\|V\|_F \le \|Y\|_F$, $\|U\|_F \le \|X\|_F$.

Claim C.1: Let $\hat{Y} = \Sigma(\Sigma + D)^{-T} Y$, then

$$\eta \triangleq 1 - \frac{\|Y\|_F}{\|\hat{Y}\|_F} \le \frac{d}{\Sigma_{\min}}, \tag{122a}$$

$$\|Y - \hat{Y}\|_F \le \frac{d}{\Sigma_{\min} - d}\|Y\|_F. \tag{122b}$$

Proof of Claim C.1: By the definition of $\hat{Y}$ we have $Y = (\Sigma + D)^T \Sigma^{-1} \hat{Y}$, then we have

$$\begin{aligned} \|Y - \hat{Y}\|_F &= \|(\Sigma + D)^T\Sigma^{-1}\hat{Y} - \hat{Y}\|_F = \|D^T\Sigma^{-1}\hat{Y}\|_F \\ &\le \|D^T\Sigma^{-1}\|_F\|\hat{Y}\|_F \le \|D^T\|_F\Sigma_{\min}^{-1}\|\hat{Y}\|_F \\ &= \frac{d}{\Sigma_{\min}}\|\hat{Y}\|_F. \end{aligned} \tag{123}$$

Using the triangular inequality and (123), we have

$$\begin{aligned} \|\hat{Y}\|_F &\le \|Y - \hat{Y}\|_F + \|Y\|_F \le \frac{d}{\Sigma_{\min}}\|\hat{Y}\|_F + \|Y\|_F, \\ &\Longrightarrow \|Y\|_F \ge (1 - \frac{d}{\Sigma_{\min}})\|\hat{Y}\|_F. \end{aligned} \tag{124}$$

The first desired inequality (122a) follows immediately from (124), and the second desired inequality (122b) is proved by combining (124) and (123). □

Combining (122a) and (121), we obtain

$$\eta \le \frac{1}{108r}. \tag{125}$$

If $\eta \le 0$, i.e. $\|\hat{Y}\|_F \le \|Y\|_F$, then $U = X, V = \hat{Y}$ already satisfy (114). From now on, we assume $\eta > 0$, i.e. $\|\hat{Y}\|_F > \|Y\|_F$. Denote $x_i^T, \hat{y}_i^T, u_i^T, v_i^T$ as the i-th row of $X, \hat{Y}, U, V$, respectively. Denote $\alpha_i \triangleq \angle(x_i, \hat{y}_i)$, i.e. the angle between the two vectors x_i and $\hat{y}_i$. Since $\langle x_i, \hat{y}_i\rangle = \Sigma_i > 0$, we have $\alpha_i \in [0, \frac{\pi}{2})$. Without loss of generality, assume

$$\alpha_1, \ldots, \alpha_s > \frac{3}{8}\pi, \quad \alpha_{s+1}, \ldots, \alpha_r \le \frac{3}{8}\pi, \tag{126}$$

where $s \in \{0, 1, \ldots, r\}$. We consider three cases and construct U, V that satisfy the desired properties in the subsequent three subsections.

$$\text{Case 1}: \sum_{i=s+1}^{r}\|\hat{y}_i\|^2 \ge \frac{2}{3}\|\hat{Y}\|_F^2, \quad \sum_{i=s+1}^{r}\|x_i\|^2 \ge \frac{2}{3}\|X\|_F^2. \tag{127a}$$

$$\text{Case 2a}: \sum_{i=1}^{s}\|\hat{y}_i\|^2 > \frac{1}{3}\|\hat{Y}\|_F^2. \tag{127b}$$

$$\text{Case 2b}: \sum_{i=1}^{s}\|x_i\|^2 > \frac{1}{3}\|X\|_F^2. \tag{127c}$$

C.3.1 Proof of Case 1

Without loss of generality, assume

$$\|x_{s+1}\| \le \|x_{s+2}\| \le \cdots \le \|x_r\|. \tag{128}$$

Let K be the smallest integer in $\{s+1, s+2, \ldots, r\}$ so that

$$\sum_{i=s+1}^{K}\|\hat{y}_i\|^2 \ge 2\sum_{j=K+1}^{r}\|\hat{y}_j\|^2. \tag{129}$$

By this definition of K, we have

$$\sum_{i=s+1}^{K-1}\|\hat{y}_i\|^2 < 2\sum_{j=K}^{r}\|\hat{y}_j\|^2. \tag{130}$$

We will shrink and extend $x_i, \hat{y}_i$ to obtain U, V. The precise definition of $U = (u_1, u_2, \ldots, u_r)^T, V = (v_1, \ldots, v_r)^T$ is given in Table VIII.

We will show that such U, V satisfy the requirements (114). The requirement (114a) follows directly from the definition of U, V and the fact $X\hat{Y}^T = \Sigma$.

We then prove the requirement (114c). We can bound $\|U - X\|_F$ as

$$\begin{aligned}
&\|U - X\|_F \\
&= \sqrt{\sum_{i>s} \|\frac{1}{1-\epsilon_i} x_i - x_i\|^2} = \sqrt{\sum_{i>s} \left(\frac{\epsilon_i}{1-\epsilon_i}\right)^2 \|x_i\|^2} \\
&\leq \frac{7\bar{\eta}}{1-7\bar{\eta}} \sqrt{\sum_{i>s} \|x_i\|^2} \leq \frac{7\bar{\eta}}{1-7\bar{\eta}} \|X\|_F \leq \frac{15}{2} \bar{\eta} \beta_T. \quad (134)
\end{aligned}$$

The bound of $\|V - \hat{Y}\|_F$ is given as

$$\begin{aligned}
\|V - \hat{Y}\|_F &= \sqrt{\sum_{i>s} \|(1-\epsilon_i)\hat{y}_i - \hat{y}_i\|^2} \\
&\leq \sqrt{\sum_{i>s} \epsilon_i^2 \|\hat{y}_i\|^2} \leq 7\bar{\eta} \|\hat{Y}\|_F.
\end{aligned}$$

Combining with the bound (123), we can bound $\|V - Y\|_F$ as

$$\begin{aligned}
&\|V - Y\|_F \leq \|V - \hat{Y}\|_F + \|\hat{Y} - Y\|_F \leq 7\bar{\eta}\|\hat{Y}\|_F + \frac{d}{\Sigma_{\min}} \|\hat{Y}\|_F \\
&= 8\bar{\eta}\|\hat{Y}\|_F \overset{(122a)}{\leq} \frac{8\bar{\eta}}{1-\bar{\eta}} \|Y\|_F \leq \frac{58}{7} \bar{\eta}\beta_T. \quad (135)
\end{aligned}$$

The first part of the requirement (114c) now follows by multiplying (134) and (135), and the second part of the requirement (114c) follows directly from (134) and (135).

At last, we prove that U, V satisfy the requirement (114b). Let

$$S_1 \triangleq \sum_{i=s+1}^{K} \|\hat{y}_i\|^2, \quad S_2 \triangleq \sum_{j=K+1}^{r} \|\hat{y}_j\|^2, \quad S_3 \triangleq \sum_{k=1}^{s} \|\hat{y}_k\|^2,$$

then (129) and (127a) imply

$$S_2 \leq S_1/2, \quad S_3 \leq (S_1 + S_2)/2 \leq 3S_1/4. \quad (136)$$

Since $\bar{\eta} = d/\Sigma_{\min} \geq \eta$, we have $(1-\eta)^2(1-\bar{\eta})^2 \geq (1-2\eta)(1-2\bar{\eta}) \geq (1-2\bar{\eta})^2$. Then

$$\begin{aligned}
&(1-\eta)^2(1-\bar{\eta})^2\|\hat{Y}\|_F^2 - \|V\|_F^2 \\
&\geq (1-2\bar{\eta})^2\|\hat{Y}\|_F^2 - \|V\|_F^2 \\
&= \sum_{i \geq s+1} ((1-2\bar{\eta})^2\|\hat{y}_i\|^2 - \|v_i\|^2) \\
&\quad + \sum_{k \leq s} ((1-2\bar{\eta})^2\|\hat{y}_k\|^2 - \|v_k\|^2) \\
&= \sum_{i \geq s+1} ((1-2\bar{\eta})^2\|\hat{y}_i\|^2 - (1-\epsilon_i)^2\|\hat{y}_i\|^2) \\
&\quad + \sum_{k \leq s} ((1-2\bar{\eta})^2\|\hat{y}_k\|^2 - \|\hat{y}_k\|^2) \\
&= \sum_{i \geq s+1} (\epsilon_i - 2\bar{\eta})(2 - \epsilon_i - 2\bar{\eta})\|\hat{y}_i\|^2 \\
&\quad - \sum_{k \leq s} 4\bar{\eta}(1-\bar{\eta})\|\hat{y}_k\|^2 \\
&\overset{(132)}{=} \sum_{s+1 \leq i \leq K} 5\bar{\eta}(2 - 5\bar{\eta} - 2\bar{\eta})\|\hat{y}_i\|^2 \\
&\quad + \sum_{K < j \leq r} (-6.5\bar{\eta})(2 + 4.5\bar{\eta} - \bar{\eta})\|\hat{y}_j\|^2 \\
&\quad - \sum_{k \leq s} 4\bar{\eta}(1-\bar{\eta})\|\hat{y}_k\|^2 \\
&= 5\bar{\eta}(2-7\bar{\eta})S_1 - 6.5\bar{\eta}(2+2.5\bar{\eta})S_2 - 4\bar{\eta}(1-\bar{\eta})S_3 \\
&\overset{(136)}{\geq} 5\bar{\eta}(2-7\bar{\eta})S_1 - 6.5\bar{\eta}(2+2.5\bar{\eta})\frac{1}{2}S_1 - 4\bar{\eta}(1-\bar{\eta})\frac{3}{4}S_1 \\
&\geq (0.5 - 41\bar{\eta})\bar{\eta}S_1 \geq 0, \quad (137)
\end{aligned}$$

where the last inequliaty follows from (121). Note that $(1-\eta)\|\hat{Y}\|_F = \|Y\|_F$, thus (137) implies

$$\|V\|_F \leq (1-\eta)(1-\bar{\eta})\|\hat{Y}\|_F = (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F,$$

which proves the second part of (114b).

We then prove the first part of (114b), i.e. $\|U\|_F \leq \|X\|_F$. Let

$$T_1 \triangleq \sum_{i=s+1}^{K} \|x_i\|^2, \quad T_2 \triangleq \sum_{j=K}^{r} \|x_j\|^2.$$

We claim that

$$T_2 \geq 2T_1. \quad (138)$$

We prove (138) by contradiction. Assume the contrary that $T_2 < 2T_1$, then $\frac{1}{3}(T_2 + T_1) < T_1$, i.e.

$$\frac{1}{3} \sum_{k=s+1}^{r} \|x_k\|^2 < \sum_{i=s+1}^{K} \|x_i\|^2 \overset{(128)}{\leq} (K-s)\|x_K\|^2. \quad (139)$$

Plugging the second inequality of (127a), i.e. $\sum_{k=s+1}^{r} \|x_k\|^2 \geq \frac{2}{3}\|X\|_F^2$, into the above relation, we obtain

$$\|X\|_F^2 \leq \frac{9}{2}(K-s)\|x_K\|^2 \leq \frac{9}{2}K\|x_K\|^2. \quad (140)$$

When $j \in \{K, K+1, \dots, r\}$, we have

$$\begin{aligned}
\Sigma_{\max} \geq \Sigma_j &= \langle x_j, \hat{y}_j \rangle = \|x_j\|\|\hat{y}_j\|\cos(\alpha_j) \\
&\overset{(128),(126)}{\geq} \|x_K\|\|\hat{y}_j\|\cos(3\pi/8).
\end{aligned}$$

which implies

$$\begin{aligned}
&\|\hat{y}_j\| \leq \omega, \text{ where } \omega \triangleq \frac{1}{\cos(3\pi/8)} \frac{\Sigma_{\max}}{\|x_K\|}, \\
&\qquad j = K, K+1, \dots, s. \quad (141)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\|\hat{Y}\|_F^2 &\overset{(127a)}{\leq} \frac{3}{2} \sum_{j=s+1}^{r} \|\hat{y}_j\|^2 \overset{(130)}{\leq} \frac{9}{2} \sum_{j=K}^{r} \|\hat{y}_j\|^2 \\
&\overset{(141)}{\leq} \frac{9}{2}(r-K+1)\omega^2. \quad (142)
\end{aligned}$$

Combining (140) and (142), and using $K(r-K+1) \leq \frac{1}{4}(r+1)^2 \leq r^2$, we get

$$\begin{aligned}
&\|X\|_F^2\|\hat{Y}\|_F^2 \leq \frac{81}{4} r^2 \|x_K\|^2 \omega^2 \\
&\overset{(141)}{=} \frac{81}{4} r^2 \|x_K\|^2 \frac{1}{\cos(3\pi/8)^2} \frac{\Sigma_{\max}^2}{\|x_K\|^2} < 140 r^2 \Sigma_{\max}^2. \quad (143)
\end{aligned}$$

According to (113), we have $\|X\|_F^2\|\hat{Y}\|_F^2 \geq \|X\|_F^2\|Y\|_F^2 \geq (\frac{3}{5})^2\beta_T^4 = \frac{9}{25}C_T^2 r^2 \Sigma_{\max}^2$; combining with (143), we get $140 > \frac{9}{25}C_T^2$, which implies $C_T^2 < 389$. This contradicts the definition (120) that $C_T = 20$, thus (138) is proved.

Now we are ready to prove the first part of (114b) as follows:

$$\begin{aligned}
&\|X\|_F^2 - \|U\|_F^2 \\
&\quad = \sum_{i\geq s+1}(\|x_i\|^2 - \|u_i\|^2) + \sum_{k\leq s}(\|x_k\|^2 - \|u_k\|^2) \\
&\quad = \sum_{i\geq s+1}(\|x_i\|^2 - \frac{1}{(1-\epsilon_i)^2}\|x_i\|^2) + 0 \\
&\quad = \sum_{i\geq s+1}\frac{\epsilon_i(\epsilon_i - 2)}{(1-\epsilon_i)^2}\|x_i\|^2 \\
&\quad = \sum_{K<j\leq r}\frac{4.5\bar{\eta}(4.5\bar{\eta}+2)}{(1+4.5\bar{\eta})^2}\|x_j\|^2 \\
&\qquad - \sum_{s+1\leq i\leq K}\frac{7\bar{\eta}(2-7\bar{\eta})}{(1-7\bar{\eta})^2}\|x_i\|^2 \\
&\quad \overset{(138)}{\geq} T_2\bar{\eta}\left[\frac{4.5(4.5\bar{\eta}+2)}{(1+4.5\bar{\eta})^2} - \frac{1}{2}\frac{7(2-7\bar{\eta})}{(1-7\bar{\eta})^2}\right] \\
&\quad \geq T_2\bar{\eta}\left[\frac{9}{(1+4.5\bar{\eta})^2} - \frac{7}{(1-7\bar{\eta})^2}\right] \geq 0,
\end{aligned}$$

where the last inequality is because $\frac{(1-7\bar{\eta})^2}{(1+4.5\bar{\eta})^2} > 0.79 > \frac{7}{9}$ when $\bar{\eta} \leq 1/(108r) < 1/100$. Thus the first part of (114b) is proved.

C.3.2 Proof of Case 2a

Denote

$$X^0 = X, Y^0 = \hat{Y}, x_k^0 = x_k, y_k^0 = \hat{y}_k, \alpha_k^0 = \alpha_k, \quad k = 1, \ldots, r. \tag{144}$$

We will define $X^i = (x_1^i, \ldots, x_r^i)^T, Y^i = (y_1^i, \ldots, y_r^i)^T$ recursively. In specific, at the i-th iteration, we will adjust X^{i-1}, Y^{i-1} to X^i, Y^i so that $\|X^i\|_F \leq \|X^{i-1}\|_F, \|Y^i\|_F < \|Y^{i-1}\|_F$ while keeping the first requirement satisfied, i.e. $X^i(Y^i)^T = \Sigma$. The angle α_k^i is defined accordingly, i.e. $\alpha_k^i \triangleq \langle x_k^i, y_k^i\rangle$.

To adjust X^{i-1}, Y^{i-1} to X^i, Y^i, we will define an operation that consists of rotation and shrinking. The basic idea is the following: since the angle between x_i^{i-1} and y_i^{i-1} is large, we can rotate x_i^{i-1} to x_i^i and shrink y_i^{i-1} to y_i^i to keep the inner product invariant, i.e. $\langle x_i^{i-1}, y_i^{i-1}\rangle = \langle x_i^i, y_i^i\rangle$. However, rotating x_i^{i-1} may destroy the orthogonal relationship between x_i^{i-1} and $y_j^{i-1}, \forall j \neq i$, thus we further rotate and shrink y_j^{i-1} to y_j^i for all $j \neq i$ so that y_j^i is orthogonal to the new vector x_i^i. Fortunately, we can prove that using such an operation we still have $\langle x_j^{i-1}, y_j^i\rangle = \Sigma_j, \forall j \neq i$.

A complete description of this operation is given in Table IX. Without loss of generality, we can make the assumption (145). In fact, if (145) does not hold, we can switch i and $m_i \triangleq \arg\min_{k\in\{i,i+1,\ldots,s\}} \alpha_k^{i-1}$ and then apply Operation 2.

We will prove that Operation 2 is valid (for D_i that is small enough), i.e. X^i, Y^i defined in Operation 2 indeed exist. The properties of X^i, Y^i obtained by Operation 2 are summarized in the following claim, which will be proved in Appendix C.4.

Claim C.2: Consider $i \in \{1, 2, \ldots, s\}$. Suppose

$$\alpha_i^{i-1} \leq \alpha_j^{i-1}, \quad \forall\, j \in \{i+1, i+2, \ldots, s\}, \tag{145}$$

and $D_i > 0$ satisfies

$$\frac{D_i}{\Sigma_i} \leq \frac{1}{12r}, \tag{146}$$

then $X^i = (x_1^i, \ldots, x_r^i)^T, Y^i = (y_1^i, \ldots, y_r^i)^T$ described in Operation 2 exist and satisfy the following properties:

$$\begin{aligned}
X^i(Y^i)^T &= \Sigma, \qquad\qquad\qquad\qquad (147\text{a})\\
\|x_k^i\| &= \|x_k^{i-1}\|, \quad \forall k,\\
\|Y^i - Y^{i-1}\|_F^2 &\leq \frac{4}{5}\frac{D_i}{\Sigma_i}(\|Y^{i-1}\|_F^2 - \|Y^i\|_F^2), \qquad (147\text{b})\\
\|X^i - X^{i-1}\|_F &= \|x_i^i - x_i^{i-1}\| \leq \frac{1}{\sqrt{3}}\frac{D_i}{\Sigma_i}\|x_i^{i-1}\|\\
\|Y^i - Y^{i-1}\|_F &\leq \frac{2}{\sqrt{3}}\frac{D_i}{\Sigma_i}\|Y^{i-1}\|_F, \qquad (147\text{c})\\
\alpha_l^i &\geq \alpha_l^{i-1} - \frac{1}{r}\frac{\pi}{24} \geq \frac{1}{3}\pi, \quad l = i, i+1, \ldots, s. \qquad (147\text{d})\\
\|y_k^{i-1}\| \geq \|y_k^i\| &\geq \|y_k^{i-1}\| - \frac{1}{10r}\|y_k^{i-1}\|, \quad k = 1, 2, \ldots, s. \qquad (147\text{e})\\
\|Y^{i-1}\|_F^2 - \|Y^i\|_F^2 &\geq \frac{5}{3}\frac{D_i}{\Sigma_i}\|y_i^i\|^2. \qquad (147\text{f})
\end{aligned}$$

We continue to prove Proposition C.1 using Claim C.2. Given any $D_1, \ldots, D_s$ that satisfy (146), we can apply a sequence of Operation 2 for $i = 1, 2, \ldots, s$ to define two sequences of matrices $Y^1, \ldots, Y^s$ and $X^1, \ldots, X^s$. Since $Y^1, \ldots, Y^s$ depend on $D_1, \ldots, D_s$, thus we can use $Y^s(D_1, \ldots, D_s)$ to denote the obtained Y^s by applying Operation 2 for $D_1, \ldots, D_s$. Obviously $Y^s(0, \ldots, 0) = Y^0$. We can also view $\|Y^s\|_F^2$ as a function of $D_1, \ldots, D_s$, denoted as

$$f(D_1, \ldots, D_s) \triangleq \|Y^s(D_1, \ldots, D_s)\|_F^2. \tag{148}$$

It can be easily seen that f is a continuous function with respect to $D_1, \ldots, D_s$.

Define[5]

$$\bar{\eta} \triangleq \frac{d}{\Sigma_{\min}} \overset{(122a)}{\geq} \eta, \quad \bar{D}_i \triangleq 9\bar{\eta}\Sigma_i, \quad i = 1, \ldots, s. \tag{149}$$

We prove that

$$f(\bar{D}_1, \ldots, \bar{D}_s) \leq (1 - 4\bar{\eta})\|\hat{Y}\|_F^2. \tag{150}$$

Suppose $\bar{X}^i, \bar{Y}^i, i = 1, \ldots, s$ are recursively defined by Operation 2 for the choices of $D_i = \bar{D}_i$ and denote $\bar{X}^0 = X, \bar{Y}^0 = \hat{Y}$. Since

$$\bar{\eta} = d/\Sigma_{\min} \overset{(121)}{\leq} 1/(108r),$$

we know that $D_i = \bar{D}_i, i = 1, \ldots, s$ as defined in (149) satisfy the condition (146), thus the property (147) holds for $\bar{X}^i, \bar{Y}^i$.

[5]In the first version of the paper, we define $\bar{D}_i \triangleq \frac{9}{2}\eta\Sigma_i \leq \frac{9}{2}\bar{\eta}\Sigma_i \leq 9\frac{d}{\Sigma_{\min}}\Sigma_i$, which is enough for proving Theorem 3.1. Here we use a slightly different definition of $\bar{D}_i$ for the purpose of proving Theorem 3.2 (linear convergence of the algorithm).

TABLE IX

OPERATION 2 THAT DEFINES X^i, Y^i, WHERE $i \in \{1, \dots, s\}$

Operation 2: Rotation and Shrinking

Input: $x_k^{i-1}, y_k^{i-1}, \alpha_k^{i-1} \triangleq \angle(x_k^{i-1}, y_k^{i-1}), k = 1, \dots, r$ and D_i.
Output: $x_k^i, y_k^i, k = 1, \dots, r$ and $\alpha_k^i \triangleq \angle(x_k^i, y_k^i)$.
Procedure:
(1) Rotate x_i^{i-1} in $\text{span}\{x_i^{i-1}, y_i^{i-1}\}$ to get x_i^i, such that $\langle x_i^i, y_i^{i-1} \rangle = \Sigma_i + D_i$.
(2) Shrink y_i^{i-1} to get y_i^i such that $\langle x_i^i, y_i^i \rangle = \Sigma_i$.
(3) For all $j \neq i$, find y_j^i in $\text{span}\{y_j^{i-1}, y_i^{i-1}\} = \text{span}_{k \neq i,j}\{x_k^{i-1}\}^\perp$ such that $y_j^i \perp x_i^i$ and $\langle x_j^{i-1}, y_j^i \rangle = \langle x_j^{i-1}, y_j^{i-1} \rangle$.
(4) Define $x_j^i \triangleq x_j^{i-1}, \forall j \neq i$.

Suppose the k-th row of $\bar{Y}^i$ is $(\bar{y}_k^i)^T$, $k = 1, \dots, r$. By (147f) and the fact $\hat{Y} = \bar{Y}^0$, we have

$$\|\hat{Y}\|_F^2 - f(\bar{D}_1, \dots, \bar{D}_s) = \|\bar{Y}^0\|_F^2 - \|\bar{Y}^s\|_F^2 = \sum_{i=1}^s (\|\bar{Y}^{i-1}\|_F^2 - \|\bar{Y}^i\|_F^2) \geq \sum_{i=1}^s \frac{5}{3} \frac{\bar{D}_i}{\Sigma_i} \|\bar{y}_i^i\|^2. \quad (151)$$

We can bound $\|\bar{y}_i^i\|$ according to (147e) as

$$\|\bar{y}_i^i\| \geq \|\bar{y}_i^{i-1}\| - \frac{1}{10r}\|\bar{y}_i^{i-1}\| \geq \|\bar{y}_i^{i-1}\| - \frac{1}{10r}\|\bar{y}_i^0\| \geq \dots \geq \|\bar{y}_i^0\| - \frac{i}{10r}\|\bar{y}_i^0\| \geq \frac{9}{10}\|\bar{y}_i^0\|.$$

Plugging into (151), we get

$$\|\hat{Y}\|_F^2 - f(\bar{D}_1, \dots, \bar{D}_s) \geq \sum_{i=1}^s \frac{5}{3} \frac{\bar{D}_i}{\Sigma_i} (\frac{9}{10})^2 \|\bar{y}_i^0\|^2 \overset{(149)}{=} 15 \frac{81}{100} \bar{\eta} \sum_{i=1}^s \|\hat{y}_i\|^2 \overset{(127b)}{>} 12\bar{\eta}\frac{1}{3}\|\hat{Y}\|_F^2 = 4\bar{\eta}\|\hat{Y}\|_F^2,$$

which immediately leads to (150).

Combining (150) and the fact $f(0, \dots, 0) = \|Y^0\|_F^2 = \|\hat{Y}\|_F^2$, we have

$$f(0, \dots, 0) = \|\hat{Y}\|_F^2 > (1 - 4\bar{\eta})\|\hat{Y}\|_F^2 = f(\bar{D}_1, \dots, \bar{D}_s).$$

Since f is continuous (in the proof of Claim C.2 in Appendix C.4, all new vectors depend continuously on D_i), and notice that $1 - 4\bar{\eta} < (1 - \bar{\eta})^4 \leq (1 - \bar{\eta})^2(1 - \eta)^2 \leq 1$, there must exist

$$0 \leq D_i \leq \bar{D}_i = 9\bar{\eta}\Sigma_i, \quad i = 1, \dots, s \quad (152)$$

such that

$$f(D_1, \dots, D_s) = (1 - \bar{\eta})^2(1 - \eta)^2\|\hat{Y}\|_F^2. \quad (153)$$

Suppose $X^i, Y^i, i = 1, \dots, s$ are recursively defined by Operation 2 for these choices of D_i, where Y^s is the simplified notation for $Y^s(D_1, \dots, D_s)$. Define

$$V \triangleq Y^s, \quad U \triangleq X^s, \quad (154)$$

By this definition of V and (148), the relation (153) can be rewritten as

$$\|V\|_F^2 = (1 - \bar{\eta})^2(1 - \eta)^2\|\hat{Y}\|_F^2. \quad (155)$$

We show that U, V defined by (154) satisfy the requirements (114). The requirement (114a) follows by the property (147a) for $i = s$. The requirement (114b) is proved as follows. Combining (155) with (122a) leads to

$$\|V\|_F = (1 - \bar{\eta})(1 - \eta)\|\hat{Y}\|_F = (1 - \bar{\eta})\|Y\|_F = (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F. \quad (156)$$

According to the property (147b), we have $\|X^i\|_F = \|X^{i-1}\|_F, i = 1, \dots, s$. Thus $\|X^s\|_F = \|X^{s-1}\|_F = \dots = \|X^0\|_F = \|X\|_F$, which implies

$$\|U\|_F = \|X\|_F. \quad (157)$$

Combining (157) and (156) leads to the requirement (114b) .

It remains to show that U, V satisfy the requirement (114c). By the property (147b), we have $\|x_k^{i-1}\| = \|x_k^i\|, \forall 1 \leq k \leq r, 1 \leq i \leq s$, which implies

$$\|x_k^i\| = \|x_k^0\| = \|x_k\|, \quad \forall 1 \leq k \leq r, \ 1 \leq i \leq s. \quad (158)$$

Note that X^i differs from X^{i-1} only in the i-th row (according to (147c)), thus

$$\|U - X\|_F = \|X^s - X^0\|_F = \sqrt{\sum_{i=1}^s \|x_i^i - x_i^{i-1}\|^2} \overset{(147c)}{\leq} \frac{1}{\sqrt{3}} \frac{D_i}{\Sigma_i} \sqrt{\sum_{i=1}^s \|x_i^{i-1}\|^2} \overset{(158)}{=} \frac{1}{\sqrt{3}} \frac{D_i}{\Sigma_i} \sqrt{\sum_{i=1}^s \|x_i\|^2} \leq \frac{1}{\sqrt{3}} \frac{D_i}{\Sigma_i} \|X\|_F \overset{(152)}{\leq} 3\sqrt{3}\bar{\eta}\|X\|_F. \quad (159)$$

Plugging $\bar{\eta} = d/\Sigma_{\min}$ and $\|X\|_F \leq \beta_T$ into the above inequality, we get

$$\|U - X\|_F \leq 3\sqrt{3}\frac{\beta_T}{\Sigma_{\min}}d. \quad (160)$$

We then bound $\|V - \hat{Y}\|_F^2$ as

$$\begin{aligned}
\|V - \hat{Y}\|_F^2 &= \|Y^s - Y^0\|_F^2 \\
&\leq s \sum_{i=1}^s \|Y^i - Y^{i-1}\|_F^2 \overset{(147b)}{\leq} s \frac{4}{5} \frac{D_i}{\Sigma_i} \sum_{i=1}^s (\|Y^{i-1}\|_F^2 - \|Y^i\|_F^2) \\
&= s \frac{4}{5} \frac{D_i}{\Sigma_i} (\|Y^0\|_F^2 - \|Y^s\|_F^2) = s \frac{4}{5} \frac{D_i}{\Sigma_i} (\|\hat{Y}\|_F^2 - \|V\|_F^2) \\
&\overset{(152)}{\leq} \frac{36}{5} s\bar{\eta} (\|\hat{Y}\|_F^2 - \|V\|_F^2) \overset{(155)}{\leq} \frac{36}{5} s\bar{\eta}(2\eta + 2\bar{\eta})\|\hat{Y}\|_F^2 \\
&\leq \frac{144}{5} r\bar{\eta}^2\|\hat{Y}\|_F^2,
\end{aligned}$$

which leads to

$$\|V - \hat{Y}\|_F \le \frac{12}{\sqrt{5}}\bar{\eta}\sqrt{r}\|\hat{Y}\|_F. \tag{161}$$

Then we can bound $\|V - Y\|_F$ as

$$\begin{aligned}\|V - Y\|_F &\le \|V - \hat{Y}\|_F + \|Y - \hat{Y}\| \\ &\overset{(161),(123)}{\le} \frac{12}{\sqrt{5}}\bar{\eta}\sqrt{r}\|\hat{Y}\|_F + \frac{d}{\Sigma_{\min}}\|\hat{Y}\|_F \\ &= (\frac{12}{\sqrt{5}} + 1)\frac{d}{\Sigma_{\min}}\sqrt{r}\|\hat{Y}\|_F \\ &\overset{(122a)}{=} (\frac{12}{\sqrt{5}} + 1)\frac{d}{\Sigma_{\min}}\sqrt{r}\|Y\|_F\frac{1}{1-\eta} \\ &< \frac{13d}{2\Sigma_{\min}}\sqrt{r}\|Y\|_F \le \frac{13\beta_T}{2\Sigma_{\min}}\sqrt{r}d,\end{aligned} \tag{162}$$

where the second last inequality is due to $(\frac{12}{\sqrt{5}} + 1)/(1 - \eta) \overset{(121)}{\le} (\frac{12}{\sqrt{5}} + 1)/(1 - \frac{1}{108}) < 6.5$. The first part of the requirement (114c) now follows by multiplying (160) and (162), and the second part of the requirement (114c) follows directly from (160) and (162).

C.3.3 Proof of Case 2b

Similar to Case 2a, denote

$$X^0 = X, Y^0 = \hat{Y}, x_k^0 = x_k, y_k^0 = \hat{y}_k, \alpha_k^0 = \alpha_k.$$

By a symmetric argument to that for Case 2a (switch the role of $U, X^j, j = 0, \ldots, s$ and $V, Y^j, j = 0, \ldots, s$), we can prove that there exist $\bar{U}, \bar{V}$ that satisfy properties analogous to (114a), (156), (157), (159) and (161), i.e.

$$\bar{U}\bar{V}^T = \Sigma, \tag{163a}$$

$$\begin{aligned}\|\bar{U}\|_F &= (1-\eta)(1-\bar{\eta})\|X^0\|_F, \\ \|\bar{V}\|_F &= \|Y^0\|_F,\end{aligned} \tag{163b}$$

$$\begin{aligned}\|\bar{V} - Y^0\|_F &\le 3\sqrt{3}\bar{\eta}\|Y^0\|_F, \\ \|\bar{U} - X^0\|_F &\le \frac{12}{\sqrt{5}}\bar{\eta}\sqrt{r}\|X^0\|_F.\end{aligned} \tag{163c}$$

We will show that the following U, V satisfy the requirements (114):

$$U \triangleq \frac{\bar{U}}{(1-\eta)(1-\bar{\eta})}, \quad V \triangleq \bar{V}(1-\eta)(1-\bar{\eta}). \tag{164}$$

The requirement (114a) follows directly from (163a) and (164). According to (163b), (164) and the facts $X^0 = X$, $\|Y^0\|_F = \|\hat{Y}\|_F = \|Y\|_F/(1-\eta)$, we have $\|U\|_F = \frac{\|\bar{U}\|_F}{(1-\eta)(1-\bar{\eta})} = \|X^0\|_F = \|X\|_F$, $\|V\|_F = \|\bar{V}\|_F(1-\eta)(1-\bar{\eta}) = \|Y^0\|_F(1-\eta)(1-\bar{\eta}) = \|Y\|_F(1-\bar{\eta})$, thus the requirement (114b) is proved.

It remains to prove the requirement (114c). We bound $\|U - X\|_F$ as

$$\begin{aligned}\|U - X\|_F &\le \|U - \bar{U}\|_F + \|\bar{U} - X\|_F \\ &\overset{(164)}{\le} 2\bar{\eta}\|U\|_F + \|\bar{U} - X^0\|_F \\ &\overset{(114b),(163c)}{\le} 2\bar{\eta}\|X\|_F + \frac{12}{\sqrt{5}}\bar{\eta}\sqrt{r}\|X^0\|_F \le \frac{15}{2}\bar{\eta}\sqrt{r}\|X\|_F \\ &\le \frac{15}{2}\frac{\beta_T}{\Sigma_{\min}}\sqrt{r}d.\end{aligned} \tag{165}$$

Using the fact $\hat{Y} = Y^0$, we bound $\|V - Y\|_F$ as

$$\begin{aligned}\|V - Y\|_F &\le \|V - \bar{V}\|_F + \|\bar{V} - \hat{Y}\|_F + \|\hat{Y} - Y\|_F \\ &\overset{(164),(123)}{\le} 2\bar{\eta}\|\bar{V}\|_F + \|\bar{V} - Y^0\|_F + \frac{d}{\Sigma_{\min}}\|\hat{Y}\|_F \\ &\overset{(163b),(163c)}{\le} 2\bar{\eta}\|\hat{Y}\|_F + 3\sqrt{3}\bar{\eta}\|Y^0\|_F + \frac{d}{\Sigma_{\min}}\|\hat{Y}\|_F \\ &= (3 + 3\sqrt{3})\frac{d}{\Sigma_{\min}}\|\hat{Y}\|_F \\ &\overset{(122a)}{=} \frac{3 + 3\sqrt{3}}{1-\eta}\frac{d}{\Sigma_{\min}}\|Y\|_F \le \frac{58\beta_T}{7\Sigma_{\min}}d.\end{aligned} \tag{166}$$

The first part of the requirement (114c) now follows by multiplying (165) and (166), and the second part follows directly from (165) and (166).

C.4 Proof of Claim 4.2

Suppose Claim C.2 holds for $1, 2, \ldots, i-1$, we prove Claim C.2 for i. By the property (147a) and (147d) of Claim C.2 for $i-1$, we have

$$X^{i-1}(Y^{i-1})^T = \Sigma. \tag{167a}$$

$$\begin{aligned}\alpha_i^{i-1} &\ge \alpha_i^{[0]} - \frac{i-1}{r}\frac{1}{24}\pi \ge \frac{3}{8}\pi \\ &\quad - \frac{1}{24}\pi + \frac{1}{24r}\pi = \frac{1}{3}\pi + \frac{1}{24r}\pi \ge \frac{1}{3}\pi.\end{aligned} \tag{167b}$$

To simplify the notations, throughout the proof of Claim C.2, we denote X^{i-1}, Y^{i-1} as X, Y and denote X^i, Y^i as X', Y'. The notations $\alpha_k^{i-1}, \alpha_k^i$ are changed accordingly to α_k, α_k'. Then (167a) and (167b) become

$$XY^T = \Sigma, \tag{168a}$$

$$\alpha_i \ge \frac{1}{3}\pi + \frac{1}{24r}\pi \ge \frac{1}{3}\pi. \tag{168b}$$

We need to prove that X', Y' exist and satisfy the properties in Claim (C.2), i.e. (with the simplification of notations)

$$X'(Y')^T = \Sigma. \tag{169a}$$

$$\begin{aligned}\|x_k'\| &= \|x_k\|, \quad \forall k, \\ \|Y' - Y\|_F^2 &\le \frac{4}{5}\frac{D_i}{\Sigma_i}(\|Y\|_F^2 - \|Y'\|_F^2).\end{aligned} \tag{169b}$$

$$\begin{aligned}\|X' - X\|_F &= \|x_i' - x_i\| \le \frac{1}{\sqrt{3}}\frac{D_i}{\Sigma_i}\|x_i\|, \\ \|Y' - Y\|_F &\le \frac{2}{\sqrt{3}}\frac{D_i}{\Sigma_i}\|Y\|_F.\end{aligned} \tag{169c}$$

$$\alpha_l' \ge \alpha_l - \frac{1}{r}\frac{\pi}{24} \ge \frac{1}{3}\pi, \quad l = i, i+1, \ldots, s. \tag{169d}$$

$$\|y_k\| \ge \|y_k'\| \ge \|y_k\| - \frac{1}{10r}\|y_k\|, \quad k = 1, 2, \ldots, s. \tag{169e}$$

$$\|Y\|_F^2 - \|Y'\|_F^2 \ge \frac{5}{3}\frac{D_i}{\Sigma_i}\|y_i\|^2. \tag{169f}$$

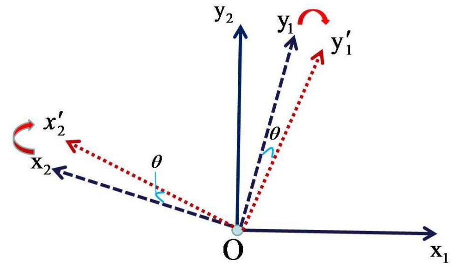

Fig. 5. A simple example to illustrate why the rotation of y_j forced by the rotation of x_i does not increase the inner product $\langle x_j, y_j\rangle$. In this example, $x_1 \perp y_2, x_2 \perp y_1$, and x_2 is rotated so that $\langle x_2, y_2\rangle$ is increased. To maintain the orthogonality, y_1 is forced to be rotated. Interestingly, $\langle y_1, x_1\rangle$ also increases.

C.4.1 Ideas of the Proof of Claim (C.2)

Before presenting the formal proof, we briefly describe its idea. The goal of Operation 2 is to reduce the norm of Y while keeping $\langle X, Y\rangle$ and $\|X\|_F$ invariant, by rotating and shrinking $x_i, y_k, k = 1, \ldots, K$ (note that $x_j, \forall j \neq i$, do no change). We first rotate x_i and shrink y_i at the same time so that the new inner product $\langle x_i', y_i'\rangle$ equals the previous one $\langle x_i, y_i\rangle$ (this step can be viewed as a combination of two steps: first rotate x_i to increase the inner product, then shrink y_i to reduce the inner product). In order to preserve the orthogonality of X and Y, we need to rotate $y_j, \forall j \neq i$, so that the new y_j' is orthogonal to x_i'.

Although the above procedure is simple, there are two questions to be answered. The first question is: will the inner product $\langle x_j, y_j\rangle$ increase as we rotate y_j, for all $j \neq i$? If yes, we could first rotate and then shrink y_j to obtain y_j' so that the new inner product $\langle x_j, y_j'\rangle$ equals $\langle x_j, y_j\rangle$, which achieves the goal of Operation 2. By resorting to the geometry (in a rigourous way) we are able to provide an affirmative answer to the above question. To gain an intuition why this is possible, we use Figure 5 to illustrate. Consider the case $i = 2$ and rotate x_2 towards y_2 to obtain x_2', then y_1 has to be rotated so that y_1' is orthogonal to x_2'. It is clear from this figure that the angle between y_1 and x_1 also decreases, or equivalently, the inner product $\langle x_1, y_1\rangle$ also increases. One might ask whether we have utilized additional assumptions on the relative positions of x_i, y_i's. In fact, we do not utilize additional assumptions; what we implicitly utilize is the fact that $\langle x_i, y_i\rangle > 0, \forall i$ (see Figure 6, Figure 7 and the paragraph after (176) for detailed explanations).

The second question is: will the angle $\alpha_j' = \angle(x_j, y_j')$ still be larger than, say, $\frac{1}{3}\pi$, for all $j > i$? If yes, then we can apply Operation 2 repeatedly for all $i = 1, 2, \ldots, s$. To provide an affirmative answer, we should guarantee that each angle decreases at most $\frac{1}{s}(\frac{3}{8}\pi - \frac{1}{3}\pi) = \frac{1}{24s}\pi$, i.e. $\angle(x_j, y_j') \geq \angle(x_j, y_j) - \frac{1}{24s}\pi, \forall\ i < j \leq s$. Unlike the first question which can be answered by reading Figure 6 and Figure 7, this question cannot be answered by just reading figures. We make some algebraic computation to obtain the following result: under the assumption that α_i is no less than α_j, during Operation 2 the amount of decrease in α_j is upper bounded by the amount of decrease in α_i, which can be further bounded above by $\frac{1}{24s}\pi$. This result explains why our proof requires the assumption $\alpha_i \geq \alpha_j, \forall\ i < j \leq s$, i.e. (145).

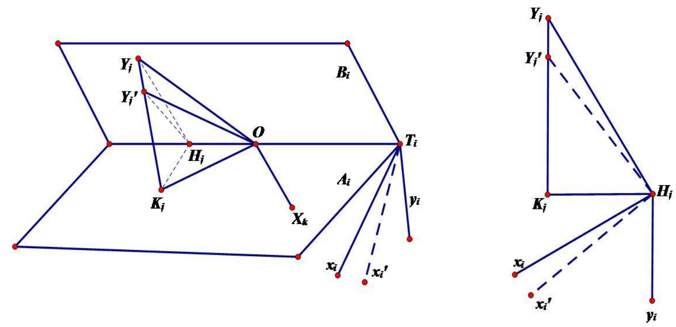

Fig. 6. Left: Space A_i, B_i, T_i, vectors x_i, y_i, x_i', x_k and some points related to y_j. Right: Some points and vectors in plane $H_j Y_j K_j = T_i^{\perp} = \text{span}\{x_i, y_i\}$. This figure shows the first possibility: x_i and K_j lie in the same side of line $H_j Y_j$.

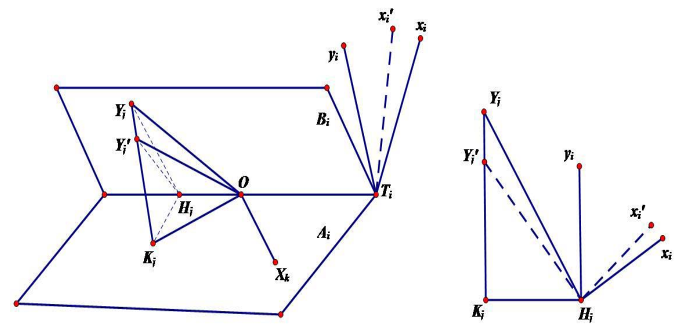

Fig. 7. Same objects as in Figure 6, but for the second possibility: x_i and K_j lie in different sides of line $H_j Y_j$.

C.4.2 Formal Proof of Claim (C.2)

We first show how to define x_i' and y_i'. Note that

$$\|x_i\|\|y_i\| = \frac{\langle x_i, y_i\rangle}{\cos\alpha_i} \geq \frac{\Sigma_i}{\cos(\frac{\pi}{3})} = 2\Sigma_i. \tag{170}$$

Since (170) implies $\frac{\Sigma_i + D_i}{\|x_i\|\|y_i\|} \leq \frac{2\Sigma_i}{\|x_i\|\|y_i\|} \leq 1$, we can define

$$\alpha_i' \triangleq \arccos(\frac{\Sigma_i + D_i}{\|x_i\|\|y_i\|}) \in [0, \frac{\pi}{2}].$$

There is a unique x_i' in the plane span$\{x_i, y_i\}$ which satisfies

$$\|x_i'\| = \|x_i\| \tag{171}$$

and $\angle(x_i', y_i) = \alpha_i'$. By the definition of α_i' above, we have

$$\langle x_i', y_i\rangle = \Sigma_i + D_i.$$

The existence of x_i' is proved. We define

$$y_i' \triangleq \frac{\Sigma_i}{\Sigma_i + D_i} y_i, \tag{172}$$

then

$$\langle x_i', y_i'\rangle = \frac{\Sigma_i}{\Sigma_i + D_i}\langle x_i', y_i\rangle = \Sigma_i. \tag{173}$$

The existence of y_i' is also proved.

Since $0 < \langle x_i, y_i\rangle = \Sigma_i < \langle x_i', y_i\rangle$, we have $\frac{\pi}{2} > \alpha_i > \alpha_i' > 0$, thus we can define

$$\theta \triangleq \alpha_i - \alpha_i' = \angle(x_i', x_i) \in (0, \alpha_i). \tag{174}$$

Fix any $j \neq i$, we then show how to define y_j'. Define

$$A_i \triangleq \text{span}_{j\neq i}\{x_j\} \perp y_i,\ \ B_i \triangleq \text{span}_{j\neq i}\{y_j\} \perp x_i,\ \ T_i \triangleq A_i \cap B_i.$$

Let $\overrightarrow{OY_j} = y_j$, $K_j \triangleq \mathcal{P}_{A_i}(Y_j), H_j \triangleq \mathcal{P}_{T_i}(Y_j)$. Then $\angle Y_j H_j K_j = \min\{\angle(x_i, y_i), \pi - \angle(x_i, y_i)\} = \angle(x_i, y_i) = \alpha_i$. Since $\alpha_i > \theta$, there exists a unique point Y_j' in the line segment $Y_j K_j$ such that

$$\angle Y_j H_j Y_j' = \theta. \tag{175}$$

Since $K_j = \mathcal{P}_{A_i}(Y_j)$ and $x_k \in A_i, \forall k \neq i$, we have $\overrightarrow{Y_j K_j} \perp x_k, \forall k \neq i$, thus

$$\overrightarrow{Y_j Y_j'} \perp x_k, \quad \forall k \neq i. \tag{176}$$

See Figure 6 and Figure 7 for the geometrical interpretation; note that T_i in general is not a line but a $r-2$ dimensional space. The righthand side subfigures represents the 2 dimensional subspace $T_i^{\perp}$; since $\text{span}\{H_j Y_j, H_j K_j\} = T_i^{\perp} = \text{span}\{x_i, y_i\}$, we can draw x_i, y_i, y_i' as the vectors starting from H_j and lying in the plane $H_j Y_j K_j = T_i^{\perp}$ in the figures. Figure 6 and Figure 7 differ in the relative position of x_i and K_j: x_i and K_j lie in the same side of line $H_j Y_j$ in Figure 6 but in different sides in Figure 7. Given the positions of x_i and H_j, Y_j, K_j, the position of y_i is determined since $y_i \perp \overrightarrow{H_j K_j}$ and $\angle(x_i, y_i) < \frac{\pi}{2}$.

In both figures, we have

$$\begin{aligned}\angle(\overrightarrow{H_j Y_j'}, x_i') &= \angle(\overrightarrow{H_j Y_j}, x_i) - \angle(x_i', x_i) + \angle Y_j H_j Y_j' \\ &\underset{(174),(175)}{=} \frac{\pi}{2} - \theta + \theta = \frac{\pi}{2}, \\ &\Longrightarrow \overrightarrow{H_j Y_j'} \perp x_i'.\end{aligned} \tag{177}$$

Now we are ready to define y_j' and establish its properties. Define

$$y_j' \triangleq \overrightarrow{OY_j'}. \tag{178}$$

Since Y_j' lies in the line segment $K_j Y_j$ and $\angle Y_j K_j O = \pi/2$, we have

$$\|y_j'\| \leq \|y_j\|. \tag{179}$$

We also have

$$y_j' = y_j + \overrightarrow{Y_j Y_j'} \in \text{span}\{y_j, y_i\} \perp x_k, \quad \forall k \neq i, j. \tag{180}$$

According to the fact $\overrightarrow{OH_j} \perp x_i'$ and (177), we have

$$y_j' = \overrightarrow{OH_j} + \overrightarrow{H_j Y_j'} \perp x_i'. \tag{181}$$

Let $k = j$ in (176), we obtain

$$0 = \langle \overrightarrow{Y_j Y_j'}, x_j \rangle = \langle y_j' - y_j, x_j \rangle = 0 \Longrightarrow \langle x_j, y_j' \rangle = \langle x_j, y_j \rangle. \tag{182}$$

We have shown that y_j' defined in (178) satisfies (180), (181) and (182), thus the existence of y_j' in Operation 2 is proved.

Having defined x_i', y_i' and $y_j', \forall j \neq i$, we further define

$$x_j' \triangleq x_j, \quad \forall j \neq i, \tag{183}$$

which completes the definition of X', Y'. In the rest, we prove that X', Y' satisfy the desired property (169).

The property (169a) can be directly proved by the definitions of X', Y'. In specific, according to (173), (182) and the definition (183), we have $\langle x_k', y_k' \rangle = \Sigma_k, \forall k$. According to the definitions (183), (172) and the fact $y_i \perp x_j, \forall j \neq i$, we have $y_i' \perp x_j', \forall j \neq i$. Together with (180) and (181), we obtain $\langle x_k', y_l' \rangle = 0, \forall k \neq l$. Thus $X'(Y')^T = \Sigma$.

Next, we prove the property (169d). We first prove

$$\alpha_i' - \alpha_i = \theta \leq \frac{1}{r} \frac{\pi}{24}. \tag{184}$$

Define $h_i \triangleq x_i' - x_i$, then

$$\|h_i\| = 2\|x_i\| \sin(\frac{\theta}{2}). \tag{185}$$

From $\langle x_i', y_i \rangle = \Sigma_i + D_i = \langle x_i, y_i \rangle + D_i$, we obtain $\langle h_i, y_i \rangle = D_i$. Note that $\langle h_i, y_i \rangle = \|h_i\| \|y_i\| \cos(\angle(h_i, y_i))$ and $\angle(h_i, y_i) = \frac{\pi}{2} - \alpha_i + \frac{\theta}{2}$, thus

$$\|h_i\| = \frac{D_i}{\|y_i\| \sin(\alpha_i - \frac{\theta}{2})}. \tag{186}$$

According to (185) and (186), we have

$$\begin{aligned}\frac{D_i}{\|x_i\| \|y_i\|} &= 2\sin(\alpha_i - \frac{\theta}{2}) \sin(\frac{\theta}{2}) \geq 2\sin(\frac{\alpha_i}{2}) \sin(\frac{\theta}{2}) \\ &\geq 2\sin(\frac{\pi}{6}) \sin(\frac{\theta}{2}) = \sin(\frac{\theta}{2}) \geq \frac{\theta}{\pi},\end{aligned}$$

where the last equality follows from the fact that $\frac{\sin(t)}{t}$ is decreasing in $t \in (0, \frac{\pi}{2}]$. Note that $\frac{D_i}{\|x_i\| \|y_i\|}$ can be upper bounded as

$$\frac{D_i}{\|x_i\| \|y_i\|} \overset{(170)}{\leq} \frac{D_i}{2\Sigma_i} \overset{(146)}{\leq} \frac{1}{24r}.$$

Combining the above two relations, we get (184).

To prove

$$\alpha_j - \alpha_j' \leq \frac{\pi}{24r}, \quad \forall j \in \{i+1, \ldots, s\}, \tag{187}$$

we only need to prove

$$\theta_j \triangleq \alpha_j - \alpha_j' \leq \theta, \quad \forall j \in \{i+1, \ldots, s\} \tag{188}$$

and then use (184). The equality (182) implies that $\|x_j\| \|y_j\| \cos(\alpha_j) = \|x_j\| \|y_j'\| \cos(\alpha_j')$, which leads to

$$\frac{\cos(\alpha_j)}{\cos(\alpha_j - \theta_j)} = \frac{\cos(\alpha_j)}{\cos(\alpha_j')} = \frac{\|y_j'\|}{\|y_j\|}.$$

For any two points P_1, P_2, we use $|P_1 P_2|$ to denote the length of the line segment $P_1 P_2$. Since $\overrightarrow{OH_j}$ is orthogonal to plane $H_j K_j Y_j$, we have

$$\frac{\|y_j'\|^2}{\|y_j\|^2} = \frac{|OH_j|^2 + |H_j Y_j'|^2}{|OH_j|^2 + |H_j Y_j|^2} \geq \frac{|H_j Y_j'|^2}{|H_j Y_j|^2},$$

where the last inequality follows from the fact that $|H_j Y_j'| \leq |H_j Y_j|$. Since $\angle Y_j H_j K_j = \alpha_i, \angle Y_j' H_j K_j = \alpha_i'$ and $\angle Y_j K_j H_j = \frac{\pi}{2}$, we have

$$\frac{|H_j Y_j'|}{|H_j Y_j|} = \frac{\sin \angle Y_j' Y_j H_j}{\sin \angle Y_j Y_j' H_j} = \frac{\sin(\pi/2 - \alpha_i)}{\sin(\pi/2 + \alpha_i')} = \frac{\cos(\alpha_i)}{\cos(\alpha_i')}.$$

According to the assumption (145) and $i < j \le s$, we have $0 \le \alpha_i \le \alpha_j \le \frac{\pi}{2}$. Since $\cos(x)/\cos(x-\theta)$ is decreasing in $[0, \frac{\pi}{2}]$, we can get

$$\frac{\cos(\alpha_i)}{\cos(\alpha_i')} = \frac{\cos(\alpha_i)}{\cos(\alpha_i - \theta)} \ge \frac{\cos(\alpha_j)}{\cos(\alpha_j - \theta)}.$$

Combining the above four relations, we get

$$\frac{\cos(\alpha_j)}{\cos(\alpha_j - \theta_j)} \ge \frac{\cos(\alpha_j)}{\cos(\alpha_j - \theta)},$$

which implies $\cos(\alpha_j - \theta) \ge \cos(\alpha_j - \theta_j)$ that immediately leads to (188). Thus we have proved (187), which combined with (184) establishes the property (169d).

Then we prove the property (169c). Since $x_j' = x_j, \forall j \neq i$, we have $\|X' - X\|_F = \|x_i' - x_i\|$, which can be bounded as

$$\begin{aligned}\|x_i' - x_i\| = \|h_i\| &\overset{(186)}{=} \frac{D_i}{\|y_i\| \sin(\alpha_i - \frac{\theta}{2})} \le \frac{\|x_i\| D_i}{\|x_i\|\|y_i\| \sin(\frac{\pi}{3})} \\ &\overset{(170)}{\le} \frac{\|x_i\| D_i}{2\Sigma_i \sin(\frac{\pi}{3})} < \frac{1}{\sqrt{3}} \frac{\|x_i\|}{\Sigma_i} D_i,\end{aligned}$$

where the first inequality is due to

$$\alpha_i - \theta/2 \ge \alpha_i - \theta \overset{(168b)}{\ge} \pi/3 + \pi/24 - \theta \overset{(184)}{\ge} \pi/3. \quad (189)$$

Thus the first part of (169c) is proved.

According to (185) and (186), we have

$$2\sin(\frac{\theta}{2}) = \frac{D_i}{\|x_i\|\|y_i\| \sin(\alpha_i - \frac{\theta}{2})} \quad (190)$$

Now we upper bound $\|y_j' - y_j\|$ as

$$\begin{aligned}\|y_j' - y_j\| &= |Y_j' Y_j| \\ &= \frac{\sin(\theta)}{\cos(\alpha_i - \theta)} |H_j Y_j| \\ &= 2\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) \frac{1}{\cos(\alpha_i - \theta)} |H_j Y_j| \\ &\overset{(190)}{=} \frac{D_i}{\|x_i\|\|y_i\| \sin(\alpha_i - \frac{\theta}{2})} \cos(\frac{\theta}{2}) \frac{1}{\cos(\alpha_i - \theta)} |H_j Y_j| \\ &\le \frac{D_i}{\|x_i\|\|y_i\| \sin(\alpha_i - \frac{\theta}{2})} \frac{1}{\cos(\alpha_i)} |H_j Y_j| \\ &\overset{(189)}{\le} \frac{D_i}{\sin(\frac{\pi}{3})\langle x_i, y_i\rangle} |H_j Y_j| \\ &\le \frac{2}{\sqrt{3}} \frac{D_i}{\Sigma_i} |H_j Y_j|, \end{aligned} \quad (191)$$

where the last inequality is due to the fact $\langle x_i, y_i \rangle = \Sigma_i$. Using $|H_j Y_j| \le \|y_j\|$, we obtain

$$\|y_j' - y_j\| \le \frac{2}{\sqrt{3}} \frac{D_i}{\Sigma_i} \|y_j\|. \quad (192)$$

According to the definition (172), we have

$$\|y_i - y_i'\| = (1 - \frac{\Sigma_i}{\Sigma_i + D_i})\|y_i\| = \frac{D_i}{\Sigma_i + D_i}\|y_i\| \le \frac{D_i}{\Sigma_i}\|y_i\|. \quad (193)$$

According to (192) (which holds for any $j \in \{1, \ldots, r\} \backslash \{i\}$) and (193), we get

$$\begin{aligned}\|Y - Y'\|_F &= \sqrt{\sum_{k=1}^r \|y_k - y_k'\|^2} \\ &\le \frac{2}{\sqrt{3}} \frac{D_i}{\Sigma_i} \sqrt{\sum_{k=1}^r \|y_k\|^2} = \frac{2}{\sqrt{3}} \frac{D_i}{\Sigma_i} \|Y\|_F,\end{aligned}$$

which proves the second part of (169c).

The property (169e) can be proved as follows. By the definition (172), we have $\|y_i'\| \le \|y_i\|$, which combined with (179) (for all $j \neq i$) leads to

$$\|y_k'\| \le \|y_k\|, \quad k = 1, \ldots, s.$$

According to (192) (for all $j \neq i$) and (193), we have $\|y_k' - y_k\| \le \frac{2}{\sqrt{3}} \frac{D_i}{\Sigma_i} \|y_k\|, \forall k$, which implies

$$\begin{aligned}\|y_k'\| \ge \|y_k\| - \|y_k' - y_k\| &\ge \|y_k\| - \frac{2}{\sqrt{3}} \frac{D_i}{\Sigma_i} \|y_k\| \\ &\overset{(146)}{\ge} \|y_k\| - \frac{1}{10r}\|y_k\|, \quad \forall k.\end{aligned}$$

Combining the above two relations we obtain the property (169e).

The property (169f) can be easily proved by (172). In fact, we have

$$\begin{aligned}&\|y_i\|^2 - \|y_i'\|^2 = (\|y_i\| - \|y_i'\|)(\|y_i\| + \|y_i'\|) \\ &\ge 2\|y_i'\|(\|y_i\| - \|y_i'\|) \overset{(172)}{=} 2\|y_i'\|(\frac{\Sigma_i + D_i}{\Sigma_i} - 1)\|y_i'\| \\ &= 2\frac{D_i}{\Sigma_i}\|y_i'\|^2 \ge 2\frac{D_i}{\Sigma_i}(\frac{11}{12})^2\|y_i\|^2 \ge \frac{5}{3}\frac{D_i}{\Sigma_i}\|y_i\|^2. \end{aligned} \quad (194)$$

where the second last inequliaty follows from $\|y_i'\| \ge \|y_i\| - \|y_i - y_i'\| \overset{(193)}{\ge} \|y_i\| - D_i\|y_i\|/\Sigma_i \overset{(146)}{\ge} 11\|y_i\|/12$. According to (179) (for all $j \neq i$), we have $\|Y\|_F^2 - \|Y'\|_F^2 \ge \|y_i\|^2 - \|y_i'\|^2$, which combined with (194) leads to the property (169f).

At last, we prove the property (169b). The first part $\|X'\|_F = \|X\|_F$ follows from (171) and (183), thus it remains to prove the second part. Denote $\varphi_j \triangleq \angle Y_j O Y_j', \beta_j \triangleq \angle Y_j O K_j$ as shown in Figure 8. Pick a point Z_j in the line segment OY_j so that $|OZ_j| = |OY_j'|$, then $|Y_j Z_j| = \|y_j\| - \|y_j'\|$. Thus we have

$$\begin{aligned}\frac{\|y_j - y_j'\|}{\|y_j\| - \|y_j'\|} &= \frac{|Y_j Y_j'|}{|Y_j Z_j|} = \frac{\sin(\angle Y_j Z_j Y_j')}{\sin(\angle Y_j Y_j' Z_j)} \\ &= \frac{\sin(\pi/2 - \varphi_j/2)}{\sin(\beta_j - \varphi_j/2)} \le \frac{1}{\sin(\beta_j - \varphi_j)}. \end{aligned} \quad (195)$$

In order to bound $1/\sin(\beta_j - \varphi_j)$,[6] we use the following bound:

$$\frac{\sin\beta_j}{\sin(\beta_j - \varphi_j)} = \frac{|Y_j K_j|}{\|y_j\|}\frac{\|y_j'\|}{|Y_j' K_j|} \le \frac{|Y_j K_j|}{|Y_j' K_j|} = \frac{\tan\alpha_i}{\tan(\alpha_i - \theta)}.$$

[6]The part from (195) to (197) can be replaced by a simpler bound $\sin(\beta_j - \varphi_j) \ge \sin(\beta_j/2) \ge \sin(\beta_j)/2$ and we can still obtain a similar bound as (199); however, by using this simpler yet looser bound, the constant coefficient $7/8$ will be replaced by a larger constant.

Fig. 8. Illustration for the proof of the property (169b).

Then we have

$$\frac{\sin\beta_j}{\sin(\beta_j-\varphi_j)}\frac{\sin(\alpha_i-\theta)}{\sin(\alpha_i)}\frac{\cos(\alpha_i-\theta)}{\cos(\alpha_i)} = \frac{\cos\alpha_i\cos\theta+\sin\alpha_i\sin\theta}{\cos(\alpha_i)} \le \frac{\sin(\theta)}{\cos(\alpha_i)}+1. \tag{196}$$

According to (190) and the fact $\cos(\alpha_i) = \langle x_i, y_i\rangle/(\|x_i\|\|y_i\|) = \Sigma_i/(\|x_i\|\|y_i\|)$, we have

$$\frac{\sin(\theta)}{\cos(\alpha_i)} \le \frac{2\sin(\theta/2)}{\cos(\alpha_i)} = \frac{D_i}{\|x_i\|\|y_i\|\sin(\alpha_i-\theta/2)}\frac{\|x_i\|\|y_i\|}{\Sigma_i} = \frac{D_i}{\Sigma_i}\frac{1}{\sin(\alpha_i-\theta/2)} \overset{(146),(189)}{\le} \frac{1}{12}\frac{1}{\sin(\pi/3)} = \frac{1}{6\sqrt{3}}.$$

Plugging the above relation into (196), we obtain

$$\frac{\sin\beta_j}{\sin(\beta_j-\varphi_j)}\frac{\sin(\alpha_i-\theta)}{\sin(\alpha_i)} \le \frac{6\sqrt{3}+1}{6\sqrt{3}}. \tag{197}$$

Combining (195) and (191), we obtain

$$\begin{aligned}&\frac{\|y_j-y_j'\|}{\|y_j\|-\|y_j'\|}\frac{\|y_j-y_j'\|}{\|y_j\|}\\ \le\ & \frac{1}{\sin(\beta_j-\varphi_j)}\frac{2}{\sqrt{3}}\frac{D_i}{\Sigma_i}\frac{|H_jY_j|}{\|y_j\|}\\ \overset{(197)}{\le}\ & \frac{2}{\sqrt{3}}\frac{D_i}{\Sigma_i}\frac{6\sqrt{3}+1}{6\sqrt{3}}\frac{|H_jY_j|}{\|y_j\|}\frac{\sin(\alpha_i)}{\sin(\beta_j)}\frac{1}{\sin(\alpha_i-\theta)}\\ =\ & \frac{6\sqrt{3}+1}{9}\frac{D_i}{\Sigma_i}\frac{1}{\sin(\alpha_i-\theta)} \overset{(189)}{\le} \frac{6\sqrt{3}+1}{9}\frac{2}{\sqrt{3}}\frac{D_i}{\Sigma_i} \le \frac{3}{2}\frac{D_i}{\Sigma_i},\end{aligned} \tag{198}$$

where the last equality is due to $|H_jY_j|\sin(\alpha_i) = |Y_jK_j| = \|y_j\|\sin(\beta_j)$.

According to (192) and (146), we obtain that $\|y_j - y_j'\| \le \frac{2}{\sqrt{3}}\frac{1}{12}\|y_j\| \le \frac{1}{8}\|y_j\|$, which further implies $\|y_j'\| + \|y_j\| \ge 2\|y_j\| - \|y_j - y_j'\| \ge \frac{15}{8}\|y_j\|$. Then by (198) we have

$$\begin{aligned}\|y_j-y_j'\|^2 &\le \frac{5\sqrt{3}+1}{6}\frac{D_i}{\Sigma_i}(\|y_j\|-\|y_j'\|)\|y_j\|\\ &\le \frac{3}{2}\frac{D_i}{\Sigma_i}(\|y_j\|-\|y_j'\|)(\|y_j'\|+\|y_j\|)\frac{8}{15}\\ &= \frac{4}{5}\frac{D_i}{\Sigma_i}(\|y_j\|^2-\|y_j'\|^2).\end{aligned} \tag{199}$$

According to the definition (172), we have

$$\begin{aligned}\frac{\|y_i\|^2-\|y_i'\|^2}{\|y_i-y_i'\|^2} &= \frac{1-(\Sigma_i)^2/(\Sigma_i+D_i)^2}{[1-\Sigma_i/(\Sigma_i+D_i)]^2}\\ &= \frac{(\Sigma_i+D_i)^2-\Sigma_i^2}{D_i^2} = \frac{D_i^2+2D_i\Sigma_i}{D_i^2} \ge 2\frac{\Sigma_i}{D_i},\end{aligned}$$

which implies

$$\|y_i-y_i'\|^2 \le \frac{1}{2}\frac{D_i}{\Sigma_i}(\|y_i\|^2-\|y_i'\|^2). \tag{200}$$

Summing up (199) for $j \in \{1,\dots,r\}\backslash\{i\}$ and (200), we obtain

$$\|Y-Y'\|_F^2 \le \frac{4}{5}\frac{D_i}{\Sigma_i}(\|Y\|_F^2-\|Y'\|_F^2),$$

which proves the second part of (169b).

D. Proofs of the Results in Section 5

D.1 Proof of Claim 5.2

The proof of this claim consists of two parts: first, by a classical result we have that M_0, the best rank-r approximation of $\frac{1}{p}\mathcal{P}_\Omega(M)$, is close to M; second, show that the scaling does not change the closeness.

We first present the following result.

Lemma D.1: Assume M is a rank r matrix of dimension $m\times n$ with $m \ge n$, and denote $M_{\max} = \|M\|_\infty$ as the maximum magnitude of the entries of M. Suppose each entry of M is included in Ω with probability $p \ge C_0\frac{\log(m+n)}{m}$, and M_0 is the best rank-r approximation of $\frac{1}{p}\mathcal{P}_\Omega(M)$. Then with probability larger than $1-1/(2n^4)$,

$$\frac{1}{mnM_{\max}^2}\|M-M_0\|_F^2 \le C_2\frac{\alpha^{\frac{3}{2}}r}{pm}, \tag{201}$$

for some numerical constant C_2.

Remark: Lemma D.1 can be found in [31]. The original version [31, Th. 1.1] holds for $M_0 = \mathrm{P}_r(\mathrm{T_r}(\mathcal{P}_\Omega(M))/p)$, where $\mathrm{T_r}(\cdot)$ denotes a trimming operator which sets to zero all rows and columns that have too many observed entries, and $\mathrm{P}_r(\cdot)$ denotes the best rank-r approximation. By standard Chernoff bound one can show that none of the rows and columns have too many observed entries with high probability, thus the conclusion of [31, Th. 1.1] holds for $M_0 = \mathrm{P}_r(\mathcal{P}_\Omega(M))/p$. The key to establish Lemma D.1 is a bound on $\|M - \frac{1}{p}\mathcal{P}_\Omega(M)\|_2$, which can be simply proved by matrix concentration inequalities; see [17, Remark 6.1.2], [4, Th. 6.3] or [7, Th. 3.5]. The proof of [31, Th. 1.1] is more complicated than applying matrix concentration inequalities since it holds for a weaker condition $|\Omega| \geq O(n)$.

Note that $\hat{X}_0, \hat{Y}_0$ defined in Table I satisfy

$$\hat{X}_0\hat{Y}_0^{\,T} = \mathrm{P}_r(\mathcal{P}_\Omega(M)/p) = M_0. \tag{202}$$

Recall that the SVD of M is $M = \hat{U}\Sigma\hat{V}$, where $\hat{U}, \hat{V}$ satisfies (12). We have

$$\begin{aligned}|M_{ij}| &= \sum_{k=1}^{r} |\hat{U}_{ik}\hat{V}_{jk}\Sigma_k| \leq \Sigma_{\max}\sum_{k=1}^{r}|\hat{U}_{ik}\hat{V}_{jk}| \\ &\leq \Sigma_{\max}\sqrt{\sum_{k=1}^{r}\hat{U}_{ik}^2}\sqrt{\sum_{k=1}^{r}\hat{V}_{jk}^2} \overset{(12)}{\leq} \Sigma_{\max}\frac{\mu r}{\sqrt{mn}}, \quad \forall i, j.\end{aligned} \tag{203}$$

The above relation implies $M_{\max} \leq \Sigma_{\max}\frac{\mu r}{\sqrt{mn}}$. Plugging this inequality and $p = |\Omega|/(mn)$ into (201), we get

$$\|M - M_0\|_F^2 \leq C_2\frac{mn\alpha^{\frac{3}{2}}r}{pm}\Sigma_{\max}^2\frac{\mu^2 r^2}{mn} = C_2 n\frac{\alpha^{\frac{3}{2}}r^3\kappa^2\mu^2}{|\Omega|}\Sigma_{\min}^2. \tag{204}$$

Plugging (202) and the assumption (27) into (204), we get

$$\hat{\delta}_0 \triangleq \|M - \hat{X}_0\hat{Y}_0^{\,T}\|_F \leq \sqrt{\frac{C_2}{C_0}}\frac{\Sigma_{\min}}{r^{1.5}\kappa^2}. \tag{205}$$

The property (a), i.e. $(X_0, Y_0) \in (\sqrt{2/3}K_1)$ follows directly from the definitions of X_0 and Y_0 in (23). We then prove the property (b), i.e. $(X_0, Y_0) \in (\sqrt{2/3}K_2)$. By (205) we have $\|M - M_0\|_F \leq \Sigma_{\min}/5 \leq \Sigma_{\max}/5$ for large enough C_0. This inequality combined with $\|M - M_0\|_F \geq \|M - M_0\|_2 \geq \|M_0\|_2 - \Sigma_{\max}$ yields

$$\|M_0\|_2 \leq \frac{6}{5}\Sigma_{\max}. \tag{206}$$

By the definitions of $\hat{X}_0, \hat{Y}_0$ (i.e. $\hat{X}_0 = \bar{X}_0 D_0^{\frac{1}{2}}$, $\hat{Y}_0 = \bar{Y}_0 D_0^{\frac{1}{2}}$, where $\bar{X}_0 D_0 \bar{Y}_0^T$ is the SVD of M_0), we have

$$\|\hat{X}_0\|_2 = \|\hat{Y}_0\|_2 = \sqrt{\|M_0\|_2} \overset{(206)}{\leq} \sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}. \tag{207}$$

Then we have

$$\|\hat{X}_0\|_F^2 \leq r\|\hat{X}_0\|_2^2 \leq \frac{6}{5}r\Sigma_{\max} \overset{(15)}{<} \frac{2}{3}\beta_T^2, \tag{208}$$

where the last inequality follows from $C_T > 9/5$. By the definition of X_0 in (23), we have $\|X_0\|_F^2 \leq \|\hat{X}_0\|_F^2 \leq \frac{2}{3}\beta_T^2$. Similarly, we can prove $\|Y_0\|_F^2 \leq \frac{2}{3}\beta_T^2$. Thus the property (b) is proved.

Next we prove the property (c), i.e. $\|M - X_0Y_0^T\|_F \leq \delta_0$. Since $\hat{X}_0, \hat{Y}_0$ satisfy $\max\{\|\hat{X}_0\|_F, \|\hat{Y}_0\|_F\} \leq \beta_T$ (due to (208) and the analogous inequality for $\hat{Y}_0$) and (205), it follows from Proposition 4.1 that there exist U_0, V_0 such that

$$U_0V_0^T = M; \tag{209a}$$

$$\|U_0\|_2 \leq \|X_0\|_2; \tag{209b}$$

$$\begin{aligned}\|U_0 - \hat{X}_0\|_F &\leq \frac{6\|\hat{Y}_0\|_2}{5\Sigma_{\min}}\hat{\delta}_0, \\ \|V_0 - \hat{Y}_0\|_F &\leq \frac{3\|\hat{X}_0\|_2}{\Sigma_{\min}}\hat{\delta}_0;\end{aligned} \tag{209c}$$

$$\begin{aligned}\|U_0^{(i)}\|^2 &\leq \frac{r\mu}{m}\beta_T^2, \\ \|V_0^{(j)}\|^2 &\leq \frac{3r\mu}{2n}\beta_T^2.\end{aligned} \tag{209d}$$

Note that the above inequalities (209b) and (209c) are not due to (48b) and (48c) of Proposition 4.1, but stronger results (99) and (107) established during the proof of Proposition 4.1.

Note that

$$\begin{aligned}\|M - X_0Y_0^T\|_F &= \|U_0(V_0 - Y_0)^T + (U_0 - X_0)Y_0^T\|_F \\ &\leq \|U_0(V_0 - Y_0)^T\|_F + \|(U_0 - X_0)Y_0^T\|_F \\ &\leq \|U_0\|_2\|V_0 - Y_0\|_F + \|U_0 - X_0\|_F\|Y_0\|_2,\end{aligned} \tag{210}$$

where the last inequality follows from Proposition B.4. Since $X_0^{(i)}$ and $\hat{X}_0^{(i)}$ has the same direction and $\|X_0^{(i)}\| \leq \|\hat{X}_0^{(i)}\|$, by Proposition B.3 we have

$$\|X_0\|_2 \leq \|\hat{X}_0\|_2 \leq \sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}. \tag{211}$$

Combining (209b) and (211), we get

$$\|U_0\|_2 \leq \sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}. \tag{212}$$

Similar to (211), we have

$$\|Y_0\|_2 \leq \sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}. \tag{213}$$

It remains to bound $\|V_0 - Y_0\|_F$ and $\|U_0 - X_0\|_F$. Let us prove the following inequality:

$$\|U_0^{(i)} - X_0^{(i)}\| \leq \|U_0^{(i)} - \hat{X}_0^{(i)}\|, \ \ \forall\, i. \tag{214}$$

If $\|\hat{X}_0^{(i)}\| \leq \sqrt{\frac{2}{3}}\beta_1$, then (214) becomes equality since $\hat{X}_0^{(i)} = X_0^{(i)}$. Thus we only need to consider the case $\|\hat{X}_0^{(i)}\| > \sqrt{\frac{2}{3}}\beta_1$. In this case by the definition of X_0 in (23) we have $\|X_0^{(i)}\| = \sqrt{\frac{2}{3}}\beta_1$. From (209d), we get

$$\|U_0^{(i)}\|^2 < \frac{3}{2}\frac{r\mu}{m}\beta_T^2 \leq \frac{2}{3}\beta_1^2 < \|\hat{X}_0^{(i)}\|^2. \tag{215}$$

For simplicity, denote $u \triangleq U_0^{(i)}, x \triangleq X_0^{(i)}, \tau \triangleq \frac{\|\hat{X}_0^{(i)}\|}{\sqrt{2/3}\beta_1} = \frac{\|\hat{X}_0^{(i)}\|}{\|x\|} > 1$. Then (215) becomes $\|u\| \leq \|x\|$ and (214) becomes $\|u - x\| \leq \|u - \tau x\|$. The latter can be transformed as follows:

$$\begin{aligned}&\|u - x\| \leq \|u - \tau x\| \\ &\quad\Longleftrightarrow \|x\|^2 - 2\langle u, x\rangle \leq \tau^2\|x\|^2 - 2\tau\langle u, x\rangle \\ &\quad\Longleftrightarrow 2(\tau - 1)\langle u, x\rangle \leq (\tau^2 - 1)\|x\|^2 \\ &\quad\Longleftrightarrow 2\langle u, x\rangle \leq (\tau + 1)\|x\|^2.\end{aligned} \tag{216}$$

Since $\langle u, x\rangle \leq \|u\|\|x\| \leq \|x\|^2$ (here we use $\|u\| \leq \|x\|$ which is equivalent to (215)) and $2 < \tau+1$, the last inequality of (216) holds, which implies that $\|u-x\| \leq \|u-\tau x\|$ holds and, consequently, (214) holds.

An immediate consequence of (214) is

$$\|U_0 - X_0\|_F \leq \|U_0 - \hat{X}_0\|_F \overset{(209c)}{\leq} \frac{5\|\hat{Y}_0\|_2}{4\Sigma_{\min}}\hat{\delta}_0 \overset{(207)}{\leq} \frac{5}{4}\sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}\frac{\hat{\delta}_0}{\Sigma_{\min}}. \tag{217}$$

Similarly, we have

$$\|V_0 - Y_0\|_F \overset{(209c)}{\leq} 3\sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}\frac{\hat{\delta}_0}{\Sigma_{\min}}. \tag{218}$$

Plugging (212), (213), (217) and (218) into (210), we get

$$\begin{aligned}
&\|M - X_0Y_0^T\|_F \\
&\leq \sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}\frac{5}{4}\sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}\frac{\hat{\delta}_0}{\Sigma_{\min}} \\
&\quad + \sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}3\sqrt{\frac{6}{5}}\sqrt{\Sigma_{\max}}\frac{\hat{\delta}_0}{\Sigma_{\min}} \\
&= (\frac{3}{2} + \frac{18}{5})\kappa\hat{\delta}_0 \\
&\overset{(205)}{\leq} \frac{51}{10}\sqrt{\frac{C_2}{C_0}\frac{\Sigma_{\min}}{r^{1.5}\kappa}} \\
&\overset{(16)}{\leq} \delta_0,
\end{aligned}$$

where the last inequality holds for $C_d \geq \frac{5}{153}\sqrt{\frac{C_0}{C_2}}$. Therefore property (c) is proved.

D.2 Proof of Claim 3.1

As mentioned in Section II-A, in this proof we only need to consider the Bernolli model that Ω includes each entry of M with probability p and the expected size S satisfies (27). Denote $d \triangleq \|M - XY^T\|_F$. Let $a = U(V-Y)^T + (U-X)V^T$, $b = (U-X)(V-Y)$, where U, V are defined with the properties in Corollary 4.1.

According to (46) we have $\|\mathcal{P}_\Omega(a)\|_F^2 \geq \frac{27}{40}pd^2$. According to (40a), we have $\|\mathcal{P}_\Omega(b)\|_F \leq \frac{1}{5}\sqrt{p}d$. Therefore, $\|\mathcal{P}_\Omega(M - XY^T)\|_F = \|\mathcal{P}_\Omega(a-b)\|_F \geq \|\mathcal{P}_\Omega(a)\|_F - \|\mathcal{P}_\Omega(b)\|_F \geq \sqrt{\frac{27}{40}}\sqrt{p}d - \frac{1}{5}\sqrt{p}d \geq \frac{3}{5}\sqrt{p}d \geq \frac{1}{\sqrt{3}}\sqrt{p}d$.

According to (40b), we have $\|b\|_F \leq \frac{1}{10}d$. According to (45) (which is a corollary of [4, Theorem 4.1]), we have $\|\mathcal{P}_\Omega(a)\|_F^2 \leq \frac{7}{6}p\|a\|_F^2 \leq \frac{7}{6}p(\|M - XY^T\|_F + \|b\|_F)^2 \leq \frac{7}{6}p(1+\frac{1}{10})^2d^2 \leq \frac{17}{12}pd^2$. Thus, $\|\mathcal{P}_\Omega(a-b)\|_F \leq \|\mathcal{P}_\Omega(a)\|_F + \|\mathcal{P}_\Omega(b)\|_F \leq (\sqrt{\frac{17}{12}} + \frac{1}{5})\sqrt{p}d \leq \sqrt{2p}d$. □

D.3 Proof of Proposition 5.1

We first provide a general condition for $(X, Y) \in K_1 \cap K_2$ (i.e. incoherent and bounded) based on the function value $\tilde{F}(X, Y)$.

Proposition D.1: Suppose the sample set Ω *satisfies* (29) *and* $\rho = 2p\delta_0^2/G_0(3/2)$, *where* δ_0 *is defined in* (16). *Suppose* (X_0, Y_0) *satisfies* (66) *and*

$$\tilde{F}(X, Y) \leq 2\tilde{F}(X_0, Y_0). \tag{219}$$

Then $(X, Y) \in K_1 \cap K_2$.

Proof of Proposition D.1: We prove by contradiction. Assume the contrary that $(X, Y) \notin K_1 \cap K_2$. By the definition of K_1, K_2 in (30), we have either $\|X^{(i)}\|^2 > \beta_1^2$ for some i, $\|Y^{(j)}\|^2 > \beta_2^2$ for some j, $\|X\|_F^2 > \beta_T^2$ or $\|Y\|_F^2 > \beta_T^2$. Hence at least one term of $G(X, Y) = \rho\sum_{i=1}^m G_0(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}) + \rho\sum_{j=1}^n G_0(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2}) + \rho G_0(\frac{3\|X\|_F^2}{2\beta_T^2}) + \rho G_0(\frac{3\|Y\|_F^2}{2\beta_T^2})$ is larger than $G_0(\frac{3}{2})$. In addition, all the other terms in the expression of $G(X, Y)$ are nonnegative, thus we have $G(X, Y) > \rho G_0(\frac{3}{2})$. Therefore,

$$\tilde{F}(X, Y) \geq G(X, Y) > \rho G_0(\frac{3}{2}) = 2p\delta_0^2. \tag{220}$$

We have

$$\begin{aligned}
\tilde{F}(X_0, Y_0) &= \frac{1}{2}\|\mathcal{P}_\Omega(M - X_0Y_0^T)\|_F^2 \\
&\leq p\|M - X_0Y_0^T\|_F^2 \leq p\delta_0^2,
\end{aligned} \tag{221}$$

where the first equality is due to $G(X_0, Y_0) = 0$ which follows from $(X_0, Y_0) \in (\sqrt{\frac{2}{3}}K_1) \cap (\sqrt{\frac{2}{3}}K_2)$, the second inequality follows from (29) and the fact $(X_0, Y_0) \in (\sqrt{\frac{2}{3}}K_1) \cap (\sqrt{\frac{2}{3}}K_2) \cap K(\delta_0) \subseteq K_1 \cap K_2 \cap K(\delta)$, and the last inequality is due to $(X_0, Y_0) \in K(\delta_0)$. Combining (220) and (221), we get

$$\tilde{F}(X, Y) > 2\tilde{F}(X_0, Y_0),$$

which contradicts (219). □

We can prove that (67) implies

$$\tilde{F}(\boldsymbol{x}_i) \leq 2\tilde{F}(\boldsymbol{x}_0), \ \forall\, i. \tag{222}$$

In fact, when (67c) holds, as the first inequality in (67c) the above relation also holds. When (67a) holds, let $\lambda = 0$ in (67a) we get (222). When (67b) holds, we have

$$\psi(\boldsymbol{x}_i, \boldsymbol{\Delta}_i; 1) \overset{(67b)}{\leq} \psi(\boldsymbol{x}_i, \boldsymbol{\Delta}_i; 0) \overset{(65b)}{=} \tilde{F}(\boldsymbol{x}_i), \tag{223}$$

which implies $\tilde{F}(\boldsymbol{x}_{i+1}) = \tilde{F}(\boldsymbol{x}_i + \boldsymbol{\Delta}_i) \overset{(65b)}{\leq} \psi(\boldsymbol{x}_i, \boldsymbol{\Delta}_i; 1) \leq \tilde{F}(\boldsymbol{x}_i)$. This relation holds for any i, thus $\tilde{F}(\boldsymbol{x}_{i+1}) \leq \tilde{F}(\boldsymbol{x}_i) \leq \cdots \leq \tilde{F}(\boldsymbol{x}_0) \leq 2\tilde{F}(\boldsymbol{x}_0)$.

Since (67) implies implies $\tilde{F}(\boldsymbol{x}_t) \leq 2\tilde{F}(\boldsymbol{x}_0)$ (see (222)), by Proposition D.1 we have $\boldsymbol{x}_t \in K_1 \cap K_2$. The rest of the proof is devoted to establish

$$\boldsymbol{x}_t \in K(\frac{2}{3}\delta), \ \forall\, t. \tag{224}$$

Define the distance of $\boldsymbol{x} = (X, Y)$ and $\boldsymbol{u} = (U, V)$ as

$$d(\boldsymbol{x}, \boldsymbol{u}) = \|XY^T - UV^T\|_F,$$

then $(X_t, Y_t) \in K(\delta) \iff \|X_tY_t^T - M\|_F \leq \delta$ can be expressed as

$$d(\boldsymbol{x}_t, \boldsymbol{u}^*) \leq \delta.$$

We first prove the following result:

Lemma D.2: If $\tilde{F}(\boldsymbol{x}) \leq 2\tilde{F}(\boldsymbol{x}_0)$, *then* $d(\boldsymbol{u}^*, \boldsymbol{x}) \notin [\frac{2}{3}\delta, \delta]$.

Proof of Lemma D.2: We prove by contradiction. Assume the contrary that

$$d(\boldsymbol{u}^*, \boldsymbol{x}) \in [\frac{2}{3}\delta, \delta]. \tag{225}$$

Since $\boldsymbol{x}_0$ satisfies (66), according to the proof of Proposition D.1 we have (221), i.e.

$$\tilde{F}(\boldsymbol{x}_0) \le p\delta_0^2. \tag{226}$$

According to Proposition D.1 and the assumption $\tilde{F}(\boldsymbol{x}) \le 2\tilde{F}(\boldsymbol{x}_0)$, we have $\boldsymbol{x} \in K_1 \cap K_2$. Together with (225) we get $\boldsymbol{x} \in K_1 \cap K_2 \cap K(\delta)$. Then we have

$$\begin{aligned}\tilde{F}(\boldsymbol{x}) &\ge \frac{1}{2}\|\mathcal{P}_\Omega(M - XY^T)\|^2 \overset{(29)}{\ge} \frac{1}{6}p\|M - XY^T\|^2 \\ &= \frac{1}{6}pd(\boldsymbol{u}^*, \boldsymbol{x})^2.\end{aligned} \tag{227}$$

Plugging $d(\boldsymbol{u}^*, \boldsymbol{x})^2 \ge (\frac{2}{3})^2\delta^2 \overset{(16)}{=} 16\delta_0^2 \overset{(226)}{\ge} 16\tilde{F}(\boldsymbol{x}_0)/p$ into (227), we get $\tilde{F}(\boldsymbol{x}) \ge \frac{8}{3}\tilde{F}(\boldsymbol{x}_0)$, which together with the assumption $\tilde{F}(\boldsymbol{x}) \le 2\tilde{F}(\boldsymbol{x}_0)$ leads to $\tilde{F}(\boldsymbol{x}) = \tilde{F}(\boldsymbol{x}_0) = 0$. Then by (227) we get $d(\boldsymbol{u}^*, \boldsymbol{x}) = 0$, which contradicts (225) since $\delta > 0$. Thus Lemma D.2 is proved.

Now we get back to the proof of (224). We prove (224) by induction on t. The basis of the induction holds due to (66) and the fact $\delta_0 = \delta/6$. Suppose $\boldsymbol{x}_t \in K(2\delta/3)$, we need to prove $\boldsymbol{x}_{t+1} \in K(2\delta/3)$. Assume the contrary that $\boldsymbol{x}_{t+1} \notin K(2\delta/3)$, i.e.

$$d(\boldsymbol{u}^*, \boldsymbol{x}_{t+1}) > \frac{2}{3}\delta. \tag{228}$$

Let $i = t+1$ in (222), we get $\tilde{F}(\boldsymbol{x}_{t+1}) \le 2\tilde{F}(\boldsymbol{x}_0)$. Then by Lemma D.2 we have

$$d(\boldsymbol{x}_{t+1}, \boldsymbol{u}^*) \notin [\frac{2}{3}\delta, \delta]; \tag{229}$$

Combining (229) and (228), we get

$$d(\boldsymbol{x}_{t+1}, \boldsymbol{u}^*) > \delta. \tag{230}$$

In the rest of the proof, we will derive a contradiction for the three cases (67a), (67b) and (67c) separately.

Case 1: (67a) holds. By the induction hypothesis, $d(\boldsymbol{x}_t, \boldsymbol{u}^*) \le \frac{2}{3}\delta$. Since $d(\boldsymbol{x}, \boldsymbol{u}^*)$ is a continuous function over $\boldsymbol{x}$, the relation $d(\boldsymbol{x}_t, \boldsymbol{u}^*) \le \frac{2}{3}\delta$ and (230) imply that there must exist some $\boldsymbol{x}' = (1-\lambda)\boldsymbol{x}_{t+1} + \lambda\boldsymbol{x}_t, \lambda \in [0, 1]$ such that

$$d(\boldsymbol{x}', \boldsymbol{u}^*) = \delta. \tag{231}$$

According to (67a), we have $\tilde{F}(\boldsymbol{x}') \le 2\tilde{F}(\boldsymbol{x}_0)$. By Lemma D.2, we have $d(\boldsymbol{u}^*, \boldsymbol{x}') \notin [\frac{2}{3}\delta, \delta]$, which contradicts (231).

Case 2: (67b) holds. Define

$$\boldsymbol{\lambda}' = \arg\min_{\lambda \in \mathbb{R}, d(\boldsymbol{x}_t + \lambda\Delta_t, \boldsymbol{u}^*) \le \delta} \psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; \lambda). \tag{232}$$

By the induction hypothesis, $d(\boldsymbol{x}_t, \boldsymbol{u}^*) \le \delta$, thus 0 lies in the feasible region of the optimization problem in (232), which implies

$$\psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; \lambda') \le \psi(\boldsymbol{x}_t, \Delta_t; 0) \overset{(65b)}{=} \tilde{F}(\boldsymbol{x}_t). \tag{233}$$

Define $\boldsymbol{x}' = \boldsymbol{x}_t + \lambda'\Delta_t$, then the feasibility of λ' for the optimization problem in (232) implies $\delta \ge d(\boldsymbol{x}', \boldsymbol{u}^*)$. Since $d(\boldsymbol{x}, \boldsymbol{u}^*)$ is a continuous function over $\boldsymbol{x}$ and $d(\boldsymbol{x}', \boldsymbol{u}^*) \le \delta \overset{(230)}{<} d(\boldsymbol{x}_{t+1}, \boldsymbol{u}^*)$, there must exist some $\boldsymbol{x}'' = (1-\epsilon)\boldsymbol{x}_{t+1} + \epsilon\boldsymbol{x}' = \boldsymbol{x}_t + (1-\epsilon+\epsilon\lambda')\boldsymbol{\Delta}_t, \epsilon \in [0, 1]$ such that

$$d(\boldsymbol{x}'', \boldsymbol{u}^*) = \delta. \tag{234}$$

Then we have

$$\begin{aligned}\tilde{F}(\boldsymbol{x}'') &\overset{(65b)}{\le} \psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; 1-\epsilon+\epsilon\lambda') \\ &\overset{(65a)}{\le} (1-\epsilon)\psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; 1) + \epsilon\psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; \lambda') \\ &\overset{(223),(233)}{\le} \tilde{F}(\boldsymbol{x}_t) \overset{(222)}{\le} 2\tilde{F}(\boldsymbol{x}_0).\end{aligned}$$

Again we apply Lemma D.2 to obtain $d(\boldsymbol{u}^*, \boldsymbol{x}'') \notin [\frac{2}{3}\delta, \delta]$, which contradicts (234).

Case 3: (67c) holds. By (66) and the fact $\delta_0 = \delta/6$ we get $d(\boldsymbol{x}_0, \boldsymbol{u}^*) \le \delta/6$. Then we have

$$d(\boldsymbol{x}_{t+1}, \boldsymbol{u}^*) \le d(\boldsymbol{x}_{t+1}, \boldsymbol{x}_0) + d(\boldsymbol{x}_0, \boldsymbol{u}^*) \overset{(67c)}{\le} \frac{5}{6}\delta + \frac{1}{6}\delta = \delta,$$

which contradicts (230).

In all three cases we have arrived at a contradiction, thus the assumption (228) does not hold, which finishes the induction step for $t+1$. Therefore, (224) holds for all t.

D.4 Proof of Claim 5.3

The sequence $\{\boldsymbol{x}_t\}$ generated by Algorithm 1 with either restricted Armijo rule or restricted line search satisfies (67c) because the sequence $\tilde{F}(\boldsymbol{x}_t)$ is decreasing and the requirement $d(\boldsymbol{x}_t, \boldsymbol{x}_0) \le 5\delta/6$ is enforced throughout computation.

Algorithm 2 and Algorithm 3 satisfy (67b) since all of them perform exact minimization of a convex upper bound of the objective function along some directions. Note that $\boldsymbol{x}_t$ should be understood as the produced solution after t "iterations" (one block of variables is updated in one "iteration"). In contrast, (X_k, Y_k) defined in these algorithms is the produced solution after k "loops" (all variables are updated once in one "loop"). For (X_k, Y_k) generated by Algorithm 2, we define $\boldsymbol{x}_{2k} = (X_k, Y_k), \boldsymbol{x}_{2k+1} = (X_{k+1}, Y_k)$ and $\psi(\boldsymbol{x}_t, \Delta_t; \lambda) = \tilde{F}(\boldsymbol{x}_t + \lambda\Delta_t)$, then ψ satisfies (65) and $\{\boldsymbol{x}_t\}_{t=0}^\infty = \{(X_k, Y_k), (X_{k+1}, Y_k)\}_{k=0}^\infty$ satisfies (67b). Similarly, for (X_k, Y_k) generated by Algorithm 3, define

$$\begin{aligned}&\boldsymbol{x}_{(m+n)k+i} \\ &\quad = (X_{k+1}^{(1)}, \ldots, X_{k+1}^{(i-1)}, X^{(i)}, X_k^{(i+1)}, \ldots, X_k^{(m)}, Y_k), \\ &\qquad\qquad i = 1, \ldots, m, \\ &\boldsymbol{x}_{(m+n)k+m+j} \\ &\quad = (X_{k+1}, Y_{k+1}^{(1)}, \ldots, Y_{k+1}^{(j-1)}, Y^{(j)}, Y_k^{(j+1)}, \ldots, Y_k^{(m)}), \\ &\qquad\qquad j = 1, \ldots, n,\end{aligned}$$

and $\psi(\boldsymbol{x}_t, \boldsymbol{\Delta}_t; \lambda) = \tilde{F}(\boldsymbol{x}_t + \lambda\boldsymbol{\Delta}_t) + \lambda_0\|\lambda\boldsymbol{\Delta}_t\|^2/2$, then ψ satisfies (65) and $\{\boldsymbol{x}_t\}_{t=0}^\infty$ satisfies (67b).

We then show that Algorithm 1 with constant stepsize $\eta < \bar{\eta}_1$ satisfies (67a) for some $\bar{\eta}_1$ when $\boldsymbol{\Omega}$ satisfies (29). We prove by induction on t. Define $\boldsymbol{x}_{-1} = \boldsymbol{x}_0$, then (67a) holds for $t = 0$. Assume (67a) holds for $t-1$, i.e., $\tilde{F}(\boldsymbol{x}_{t-1} + \lambda\boldsymbol{\Delta}_{t-1}) \le 2\tilde{F}(\boldsymbol{x}_0), \forall\lambda \in [0, 1]$, where $\boldsymbol{\Delta}_t = \boldsymbol{x}_t - \boldsymbol{x}_{t-1}$. In particular, we have $\tilde{F}(\boldsymbol{x}_t) \le 2\tilde{F}(\boldsymbol{x}_0)$, which together with the assumption that $\boldsymbol{\Omega}$ satisfies (29) leads to (by Proposition (D.1))

$$\boldsymbol{x}_t \in K_1 \cap K_2.$$

Thus $\max\{\|X_t\|_F, \|Y_t\|_F\} \leq \beta_T$, $\|X_t^{(i)}\| \leq \beta_1, \forall i$, and $\|Y_t^{(j)}\| \leq \beta_2, \forall j$. Then we have

$$\begin{aligned}
&\|\nabla_X \tilde{F}(\boldsymbol{x}_t)\|_F \\
&= \|\nabla_X F(\boldsymbol{x}_t) + \nabla_X G(\boldsymbol{x}_t)\|_F \\
&\leq \|\mathcal{P}_\Omega(X_t Y_t^T - M)Y_t\|_F + \left\|\rho \sum_{i=1}^m G_0'\Big(\frac{3\|X_t^{(i)}\|^2}{2\beta_1^2}\Big)\frac{3\bar{X}_t^{(i)}}{\beta_1^2}\right\|_F \\
&\quad + \left\|\rho G_0'\Big(\frac{3\|X_t\|_F^2}{2\beta_T^2}\Big)\frac{3X_t}{\beta_T^2}\right\|_F \\
&\leq \|\mathcal{P}_\Omega(X_t Y_t^T - M)\|_F\|Y_t\|_F + \frac{3\rho\|X_t\|_F}{\beta_1^2} + \frac{3\rho\|X_t\|_F}{\beta_T^2} \\
&\leq \sqrt{\tilde{F}(\boldsymbol{x}_t)}\beta_T + \frac{6\rho\|X_t\|_F}{\beta_1^2} \\
&\leq \sqrt{2\tilde{F}(\boldsymbol{x}_0)}\beta_T + \frac{6\rho\beta_T}{\beta_1^2},
\end{aligned}$$

where in the second inequality we use $G_0'(\frac{3\|X_t^{(i)}\|^2}{2\beta_1^2}) \leq G_0'(\frac{3}{2}) = 1$ and $G_0'(\frac{3\|X\|_F^2}{2\beta_T^2}) \leq G_0'(\frac{3}{2}) = 1$. Assume

$$\bar{\eta}_1 \leq \frac{1}{4\beta_T^2}. \tag{235}$$

Recall that $\eta \leq \bar{\eta}_1$, thus we have

$$\begin{aligned}
\|X_{t+1}\|_F &\leq \|X_t\|_F + \eta\|\nabla_X \tilde{F}(\boldsymbol{x}_t)\|_F \\
&\leq \beta_T + \frac{1}{4\beta_T^2}\left(\sqrt{2\tilde{F}(\boldsymbol{x}_0)}\beta_T + \frac{6\rho\beta_T}{\beta_1^2}\right) \\
&\overset{(221)}{\leq} \beta_T + \frac{1}{4\beta_T}\left(\sqrt{2p}\delta_0 + \frac{6\rho}{\beta_1^2}\right) \triangleq c_1.
\end{aligned} \tag{236}$$

By a similar argument, we can prove $\|Y_{t+1}\|_F \leq c_1$, thus $\boldsymbol{x}_{t+1} = (X_{t+1}, Y_{t+1}) \in \Gamma(c_1)$ (recall the definition of $\Gamma(\cdot)$ in (20) is $\Gamma(\beta) = \{(X,Y) \mid \|X\|_F \leq \beta, \|Y\|_F \leq \beta\}$). Since $(X_t, Y_t) \in \Gamma(\beta_T) \subseteq \Gamma(c_1)$ and $\Gamma(c_1)$ is a convex set, we have that the line segment connecting $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$, denoted as $[\boldsymbol{x}_t, \boldsymbol{x}_{t+1}]$, lies in $\Gamma(c_1)$. Then by Claim 2.1 we have that $\nabla\tilde{F}$ is Lipschitz continuous in $[\boldsymbol{x}_t, \boldsymbol{x}_{t+1}]$ with Lipschitz constant

$$L_1 = L(c_1) = 4c_1^2 + 54\rho\frac{c_1^2}{\beta_1^4} \geq L(\beta_T) \geq 4\beta_T^2, \tag{237}$$

where the last inequality is due to the fact $c_1 \geq \beta_T$. Define (note c_1 is defined by (236))

$$\bar{\eta}_1 \triangleq \frac{1}{L_1} = \frac{1}{4c_1^2 + 54\rho\frac{c_1^2}{\beta_1^4}}, \tag{238}$$

then $\bar{\eta}_1 \leq \frac{1}{L(\beta_T)} \leq \frac{1}{4\beta_T^2} = \frac{1}{4\beta_T^2}$, which is consistent with (235).

It follows from a classical descent lemma (see, e.g., [50, Proposition A.24]) that

$$\begin{aligned}
&\tilde{F}(\boldsymbol{x}_t - \lambda\eta\nabla\tilde{F}(\boldsymbol{x}_t)) \\
&\quad\leq \tilde{F}(\boldsymbol{x}_t) - \langle\lambda\eta\nabla\tilde{F}(\boldsymbol{x}_t), \nabla\tilde{F}(\boldsymbol{x}_t)\rangle + \frac{L_1}{2}\|\boldsymbol{\lambda}\eta\nabla\tilde{F}(\boldsymbol{x}_t)\|^2 \\
&\quad= \tilde{F}(\boldsymbol{x}_t) + \|\nabla\tilde{F}(\boldsymbol{x}_t)\|^2\Big(\frac{L_1}{2}\lambda^2\eta^2 - \lambda\eta\Big) \\
&\quad\leq \tilde{F}(\boldsymbol{x}_t) - \frac{\lambda\eta}{2}\|\nabla\tilde{F}(\boldsymbol{x}_t)\|^2 \\
&\quad\leq \tilde{F}(\boldsymbol{x}_t) \\
&\quad\leq 2\tilde{F}(\boldsymbol{x}_0), \quad \forall\lambda \in [0,1],
\end{aligned} \tag{239}$$

where the second inequality follows from the fact that $\lambda\eta \leq \eta \leq \bar{\eta}_1 = 1/L_1$. This finishes the induction step (note that $\boldsymbol{\Delta}_t = \boldsymbol{x}_{t+1} - \boldsymbol{x}_t = -\eta\nabla\tilde{F}(\boldsymbol{x}_t)$), thus (67a) is proved.

Finally, we show that Algorithm 4 (SGD) satisfies (67a) with $\boldsymbol{x}_t = (X_k, Y_k)$ representing the produced solution after the t-th loop, provided that $\boldsymbol{\Omega}$ satisfies (29). Denote $N = |\Omega| + m + n + 2$ and $\boldsymbol{x}_{k,i} = (X_{k,i}, Y_{k,i}), i = 1, \ldots, N$. We prove (67a) by induction on t. Define $\boldsymbol{x}_{-1} = \boldsymbol{x}_0$, then (67a) holds for $t = 0$. Assume (67a) holds for $0, 1, \ldots t-1$, i.e., $\tilde{F}(\boldsymbol{x}_k + \lambda\boldsymbol{\Delta}_k) \leq 2\tilde{F}(\boldsymbol{x}_0), \forall\lambda \in [0,1]$, where $\boldsymbol{\Delta}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$, $0 \leq k \leq t-1$. In particular, we have $\tilde{F}(\boldsymbol{x}_t) \leq 2\tilde{F}(\boldsymbol{x}_0)$, which together with the assumption that $\boldsymbol{\Omega}$ satisfies (29) leads to (by Proposition (D.1))

$$\boldsymbol{x}_t \in K_1 \cap K_2. \tag{240}$$

Now we show that there exist constants $c_{1,i}, c_{2,i}$, $i = 0, 1, \ldots, N$ (independent of t) so that

$$\max\{\|X_{t,i}\|_F, \|Y_{t,i}\|_F\} \leq c_{1,i}, \tag{241a}$$

$$\max\{\|\nabla_X f_{i+1}(\boldsymbol{x}_{t,i})\|_F, \|\nabla_Y f_{i+1}(\boldsymbol{x}_{t,i-1})\|_F\} \leq c_{2,i}. \tag{241b}$$

We prove (241) by induction on i. When $i = 0$, since by (240) we have $\max\{\|X_{t,0}\|_F, \|Y_{t,0}\|_F\} = \max\{\|X_t\|_F, \|Y_t\|_F\} \leq \beta_T$, thus (241a) holds for $c_{1,0} = \beta_T$.

Suppose (241a) holds for i, we prove (241b) holds for i with suitably chosen $c_{2,i}$. Note that f_{i+1} can be one of the five different functions in (26). When f_{i+1} equals some F_{jl}, we have

$$\begin{aligned}
&\|\nabla_X f_{i+1}(\boldsymbol{x}_{t,i})\|_F \\
&\quad= \|\nabla_X F_{j,l}(\boldsymbol{x}_{t,i})\|_F = |(X_{t,i}^{(j)})^T Y_{t,i}^{(l)} - M_{jl}|\|Y_{t,i}^{(l)}\| \\
&\quad\leq (\|X_{t,i}\|_F\|Y_{t,i}\|_F + M_{\max})\|Y_{t,i}\|_F \leq (c_{1,i}^2 + M_{\max})c_{1,i}.
\end{aligned}$$

When $f_{i+1}(X,Y)$ equals some $G_{1j}(X)$, we have (see (24) for the expression of $\nabla_X G_{1j}$)

$$\begin{aligned}
\|\nabla_X f_{i+1}(\boldsymbol{x}_{t,i})\|_F &= \|\nabla_X G_{1j}(X_{t,i})\|_F \\
&= \rho G_0'\Big(\frac{3\|X_{t,i}^{(j)}\|^2}{2\beta_1^2}\Big)\frac{3\|X_{t,i}^{(j)}\|}{\beta_1^2} \\
&\leq \rho G_0'\Big(\frac{3c_{1,i}^2}{2\beta_1^2}\Big)\frac{3c_{1,i}}{\beta_1^2} \leq \rho G_0'\Big(\frac{3c_{1,i}^2}{2\beta_T^2}\Big)\frac{3c_{1,i}}{\beta_T^2}.
\end{aligned}$$

When $f_{i+1}(X, Y)$ equals some $G_3(X)$, we have

$$\begin{aligned}\|\nabla_X f_{i+1}(\boldsymbol{x}_{t,i})\|_F &= \|\nabla_X G_3(X_{t,i})\|_F \\ &= \rho G_0'(\frac{3\|X_{t,i}\|_F^2}{2\beta_T^2})\frac{3\|X_{t,i}\|_F}{\beta_T^2} \\ &\leq \rho G_0'(\frac{3c_{1,i}^2}{2\beta_T^2})\frac{3c_{1,i}}{\beta_T^2}.\end{aligned}$$

When $f_{i+1}(X, Y)$ equals some $G_{2j}(Y)$ or $G_4(Y)$ that only depend on Y, we have $\nabla_X f_{i+1}(\boldsymbol{x}_{t,i}) = 0$. Let

$$c_{2,i} \triangleq \max\left\{(c_{1,i}^2 + M_{\max})c_{1,i},\ \ \rho G_0'(\frac{3c_{1,i}^2}{2\beta_T^2})\frac{3c_{1,i}}{\beta_T^2}\right\},$$

then no matter what kind of function f_{i+1} is, we always have $\|\nabla_X f_{i+1}(\boldsymbol{x}_{t,i})\|_F \leq c_{2,i}$. Similarly, $\|\nabla_Y f_{i+1}(\boldsymbol{x}_{t,i})\|_F \leq c_{2,i}$. Thus (241b) holds for i.

Suppose (241b) holds for $i-1$, we prove that (241a) holds for i with suitably chosen $c_{1,i}$. In fact,

$$\begin{aligned}\|X_{t,i}\|_F &= \|X_{t,i-1} - \eta_t \nabla_X f_i(\boldsymbol{x}_{t,i-1})\|_F \\ &\leq \|X_{t,i-1}\|_F + \eta_t\|\nabla_X f_i(\boldsymbol{x}_{t,i-1})\|_F \\ &\leq c_{1,i-1} + \bar{\eta}c_{2,i-1},\end{aligned}$$

thus (241a) holds for $c_{1,i} = c_{1,i-1} + \bar{\eta}c_{2,i-1}$. This finishes the induction proof of (241).

In Claim 2.1, we have proved that $\nabla\tilde{F}$ is Lipschitz continuous with Lipschitz constant $L(\beta_0) = 4\beta_0 + 54\rho\frac{\beta_0^2}{\beta_1^4}$ in the set $\Gamma(\beta_0)$ (the definition of $\Gamma(\cdot)$ is given in (20)). By a similar argument (or set irrelevant rows of X, Y, U, V to zero in the proof of Claim (2.1)), we can prove that each ∇f_i is also Lipschitz continuous with Lipschitz constant $L(\beta_0) = 4\beta_0 + 54\rho\frac{\beta_0^2}{\beta_1^4}$ in the set $\Gamma(\beta_0)$. Then we have

$$\|\nabla f_i(\boldsymbol{x}_{t,i-1}) - \nabla f_i(\boldsymbol{x}_t)\|_F \leq c_{i-1}'\|\boldsymbol{x}_{t,i-1} - \boldsymbol{x}_t\|_F, \quad i = 1, \ldots, N, \tag{242}$$

where $c_{i-1}' = L(c_{1,i-1})$.

Note that $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t + \sum_{i=1}^N(\boldsymbol{x}_{t,i} - \boldsymbol{x}_{t,i-1}) = \boldsymbol{x}_t - \eta_t\sum_{i=1}^N \nabla f_i(\boldsymbol{x}_{t,i-1})$. We can express SGD as an approximate gradient descent method:

$$\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - \eta_t(\nabla\tilde{F}(\boldsymbol{x}_t) + w_t), \tag{243}$$

where the error

$$\begin{aligned}w_t &= \sum_{i=1}^N \nabla f_i(\boldsymbol{x}_{t,i-1}) - \nabla\tilde{F}(\boldsymbol{x}_t) \\ &= \sum_{i=1}^N(\nabla f_i(\boldsymbol{x}_{t,i-1}) - \nabla f_i(\boldsymbol{x}_t)).\end{aligned}$$

Following the analysis in [61, Lemma 1], we can bound each term $\nabla f_i(\boldsymbol{x}_{t,i-1}) - \nabla f_i(\boldsymbol{x}_t)$ as

$$\begin{aligned}&\|\nabla f_i(\boldsymbol{x}_{t,i-1}) - \nabla f_i(\boldsymbol{x}_t)\|_F \overset{(242)}{\leq} c_{i-1}'\|\boldsymbol{x}_{t,i-1} - \boldsymbol{x}_t\|_F \\ &= \eta_t c_{i-1}'\|\sum_{l=1}^{i-1}\nabla f_l(\boldsymbol{x}_{t,l-1})\|_F \overset{(241b)}{\leq} \eta_t c_{i-1}'\sum_{l=1}^{i-1}\sqrt{2}c_{2,l}.\end{aligned}$$

Plugging this inequality for $i = 1, \ldots, N$ into the expression of w_t, we obtain an upper bound of the error w_t:

$$\|w_t\|_F \leq \eta_t c_0, \tag{244}$$

where $c_0 \triangleq \sum_{i=1}^N(c_{i-1}'\sum_{l=1}^{i-1}\sqrt{2}c_{2,l})$ is a constant.

Applying (241a) for $i = N$, we get $\max\{\|X_{t+1}\|_F, \|Y_{t+1}\|_F\} \leq c_{1,N}$, thus $\boldsymbol{x}_{t+1} \in \Gamma(c_{1,N})$. Since $\boldsymbol{x}_t \in \Gamma(\beta_T) \subseteq \Gamma(c_{1,N})$ and $\Gamma(c_{1,N})$ is a convex set, we have that the line segment connecting $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ lies in $\Gamma(c_{1,N})$. Then by Claim 2.1 we have that $\nabla\tilde{F}$ is Lipschitz continuous over this line segment with Lipschitz constant $L' = L(c_{1,N})$. It follows from a classical descent lemma (see, e.g., [50, Proposition A.24]) that

$$\tilde{F}(\boldsymbol{x}_{t+1}) \leq \tilde{F}(\boldsymbol{x}_t) + \langle\boldsymbol{x}_{t+1} - \boldsymbol{x}_t, \nabla\tilde{F}(\boldsymbol{x}_t)\rangle + \frac{L'}{2}\|\boldsymbol{x}_{t+1} - \boldsymbol{x}_t\|_F^2.$$

Using the expression (243), the above relation becomes

$$\begin{aligned}\tilde{F}(\boldsymbol{x}_{t+1}) - \tilde{F}(\boldsymbol{x}_t) &\leq -\eta_t\langle\nabla\tilde{F}(\boldsymbol{x}_t) + w_t, \nabla\tilde{F}(\boldsymbol{x}_t)\rangle \\ &\quad + \frac{L'}{2}\eta_t^2\|\nabla\tilde{F}(\boldsymbol{x}_t) + w_t\|_F^2.\end{aligned} \tag{245}$$

Plugging

$$\begin{aligned}&-\eta_t\langle w_t, \nabla\tilde{F}(\boldsymbol{x}_t)\rangle \leq \eta_t\|w_t\|_F\|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F \\ &\overset{(244)}{\leq} \eta_t^2 c_0\|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F \leq \frac{1}{2}\eta_t^2 c_0(1 + \|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2)\end{aligned}$$

and

$$\begin{aligned}&\frac{1}{2}\|\nabla\tilde{F}(\boldsymbol{x}_t) + w_t\|_F^2 \leq \|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2 + \|w_t\|_F^2 \\ &\overset{(244)}{\leq} \|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2 + \eta_t^2 c_0^2\end{aligned}$$

into (245), we get

$$\begin{aligned}&\tilde{F}(\boldsymbol{x}_{t+1}) - \tilde{F}(\boldsymbol{x}_t) \\ &\leq -\eta_t\|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2 + \frac{1}{2}\eta_t^2 c_0(1 + \|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2) \\ &\quad + L'\eta_t^2(\|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2 + \eta_t^2 c_0^2) \\ &= (\frac{1}{2}\eta_t^2 c_0 + \eta_t^2 L' - \eta_t)\|\nabla\tilde{F}(\boldsymbol{x}_t)\|_F^2 + \eta_t^2(\frac{1}{2}c_0 + L'\eta_t^2 c_0^2).\end{aligned} \tag{246}$$

Pick

$$\bar{\eta} \triangleq \frac{1}{c_0 + 2L'}.$$

Since $\eta_t \leq \bar{\eta}$, we have $\frac{1}{2}\eta_t^2 c_0 + \eta_t^2 L' - \eta_t \leq -\eta_t/2$ and $L'\eta_t^2 c_0^2 \leq L'c_0^2\frac{1}{(c_0+2L')^2} \leq \frac{c_0}{8}$ (the last inequality follows from $(c_0 + 2L')^2 \geq 8c_0 L'$). Plugging these two inequalities into (246), we obtain

$$\tilde{F}(\boldsymbol{x}_{t+1}) - \tilde{F}(\boldsymbol{x}_t) \leq \eta_t^2 c_0.$$

By the same argument we can prove

$$\tilde{F}(\boldsymbol{x}_{k+1}) - \tilde{F}(\boldsymbol{x}_k) \leq \eta_k^2 c_0, \quad k = 0, 1, \ldots, t.$$

Summing up these inequalities, we get

$$\tilde{F}(\boldsymbol{x}_{t+1}) \leq \tilde{F}(\boldsymbol{x}_0) + \sum_{k=0}^t \eta_k^2 c_0 \leq \tilde{F}(\boldsymbol{x}_0) + \eta_{\text{sum}}c_0.$$

where the last inequality follows from the assumption $\sum_{k=0}^{\infty} \eta_k^2 \leq \eta_{\text{sum}}$. Pick

$$\eta_{\text{sum}} \triangleq \frac{\tilde{F}(\boldsymbol{x}_0)}{c_0},$$

the above relation becomes

$$\tilde{F}(\boldsymbol{x}_{t+1}) \leq 2\tilde{F}(\boldsymbol{x}_0).$$

By a similar argument, we can prove

$$\tilde{F}(\boldsymbol{x}_t + \lambda(\boldsymbol{x}_{t+1} - \boldsymbol{x}_t)) \leq 2\tilde{F}(\boldsymbol{x}_0), \ \forall \ \lambda \in [0, 1],$$

which completes the induction. Thus we have proved that Algorithm 4 (SGD) satisfies (67a) with suitably chosen $\bar{\eta}$ and η_{sum}.

D.5 Proof of Claim 5.1

For Algorithm 1 with constant stepsize $\eta < \bar{\eta}_1$ (defined in (238)), since the objective value $\tilde{F}(\boldsymbol{x}_t)$ is decreasing, we have $\tilde{F}(\boldsymbol{x}_t) \leq \tilde{F}(\boldsymbol{x}_0)$. By Proposition D.1 this implies that the algorithm generates a sequence in $K_1 \cap K_2$. By Claim 2.1 and the fact $K_2 = \Gamma(\beta_T)$ (see the definitions of K_2 in (30) and the definition of $\Gamma(\cdot)$ in (20)), $\nabla \tilde{F}$ is Lipschitz continuous with Lipschitz constant $L(\beta_T)$ over the set K_2. According to [50, Proposition 1.2.3], each limit point of the sequence generated by Algorithm 1 with constant stepsize $\eta < \bar{\eta}_1 \overset{(238)}{\leq} 2/L(\beta_T)$ is a stationary point of problem (P1).

We then consider Algorithm 1 with stepsize chosen by the restricted Armijo rule. The proof of [50, Proposition 1.2.1] for the standard Armijo rule can not be directly applied, and some extra effort is needed. For the restricted Armijo rule, the procedure of picking the stepsize η_k can be viewed as a two-phase approach. In the first phase, we find the smallest nonnegative integer so that the distance requirement is fulfilled, i.e.

$$i_1 \triangleq \min\{i \in \mathbb{Z}^+ \mid d(\boldsymbol{x}_k(\xi^i s_0), \boldsymbol{x}_0) \leq \frac{5}{6}\delta\}, \tag{247}$$

where $\mathbb{Z}^+$ denotes the set of nonnegative integers, and let $\bar{s}_k = \xi^{i_1} s_0$. Since

$$d(\boldsymbol{x}_k(0), s_0) = d(\boldsymbol{x}_{k-1}, \boldsymbol{x}_0) \leq \frac{2}{3}\delta, \tag{248}$$

(according to Proposition 5.1 and Claim 5.3), such an integer i_1 must exist. In the second phase, find the smallest nonnegative integer so that the reduction requirement is fulfilled, i.e.

$$\begin{aligned} i_2 \triangleq \min\{i \in \mathbb{Z}^+ \mid & \tilde{F}(\boldsymbol{x}_k(\xi^i \bar{s}_k)) \\ & \leq \tilde{F}(\boldsymbol{x}_{k-1}) - \sigma \xi^i \bar{s}_k \|\nabla \tilde{F}(\boldsymbol{x}_{k-1})\|_F^2\}, \end{aligned} \tag{249}$$

and let $\eta_k = \xi^{i_2} \bar{s}_k = \xi^{i_1 + i_2} s_0$.

Note that the second phase follows the same procedure as the standard Armijo rule (see [50, eq. (1.11)]). Hence the difference between the standard Armijo rule and the restricted Armijo rule can be viewed as the following: in each iteration the former starts from a fixed initial stepsize s while the latter starts from a varying initial stepsize $\bar{s}_k$. We notice that the proof of [50, Proposition 1.2.1] does not require the initial stepsizes to be constant, but rather the following property: if the final stepsize η_k goes to zero for a subsequence $k \in \mathcal{K}$, then for large enough $k \in \mathcal{K}$ the initial stepsize must be reduced at least once (see the remark after [50, eq. (1.17)]). This property also holds when the initial stepsize is lower bounded (asymptotically). In the following, we will prove that for the restricted Armijo rule the initial stepsize $\bar{s}_k$ is lower bounded (asymptotically), and then show how to apply the proof of [50, Proposition 1.2.1] to the restricted Armijo rule.

We first prove that the sequence $\{\bar{s}_k\}$ is lower bounded (asymptotically), i.e.

$$\liminf_{k \to \infty} \bar{s}_k > 0. \tag{250}$$

Assume the contrary that $\liminf_{k \to \infty} \bar{s}_k = 0$, i.e. there exists a subsequence $\{\bar{s}_k\}_{k \in \mathcal{K}}$ that converges to zero. Since s_0 is a fixed scalar, we can assume $\bar{s}_k < s_0, \forall k \in \mathcal{K}$, thus the corresponding $i_1 > 0$ for all $k \in \mathcal{K}$. By the definition of i_1 in (247), we know that $i_1 - 1$ does not satisfy the distance requirement; in other words, we have

$$d(\boldsymbol{x}_k(\xi^{-1}\bar{s}_k), \boldsymbol{x}_0) > \frac{5}{6}\delta.$$

Denote $g_{k-1} \triangleq \nabla \tilde{F}(\boldsymbol{x}_{k-1})$, then the above relation becomes

$$\begin{aligned} \frac{5}{6}\delta &< d(\boldsymbol{x}_{k-1} - \xi^{-1}\bar{s}_k g_{k-1}, \boldsymbol{x}_0) \\ &\leq d(\boldsymbol{x}_{k-1}, x_0) + \xi^{-1}\bar{s}_k \|g_{k-1}\|_F \\ &\overset{(248)}{\leq} \frac{2}{3}\delta + \xi^{-1}\bar{s}_k \|g_{k-1}\|_F, \end{aligned}$$

implying

$$\frac{1}{6}\xi\delta \leq \bar{s}_k \|g_{k-1}\|_F.$$

Since $\frac{1}{6}\xi\delta$ is a constant and $\{\bar{s}_k\}_{k \in \mathcal{K}}$ converges to zero, the above relation implies that $\{\|g_{k-1}\|_F\}_{k \in \mathcal{K}}$ goes to infinity. However, it is easy to verify that $\|g_{k-1}\|_F = \|\nabla \tilde{F}(\boldsymbol{x}_{k-1})\|_F$ is bounded above by a universal constant when $\|\boldsymbol{x}_{k-1}\|_F \leq \beta_T$ (note that $\|\boldsymbol{x}_{k-1}\|_F \leq \beta_T$ holds due to Proposition 5.1 and Claim 5.3)), which is a contradiction. Therefore, (250) is proved.

Now we prove that each limit point of the sequence $\{\boldsymbol{x}_k\}$ generated by Algorithm 1 with restricted Armijo rule is a stationary point. Assume the contrary that there exists a limit point $\bar{\boldsymbol{x}}$ with $\nabla \tilde{F}(\bar{\boldsymbol{x}}) \neq 0$, and suppose the subsequence $\{\boldsymbol{x}_k\}_{k \in \mathcal{K}}$ converges to $\bar{\boldsymbol{x}}$. By the same argument as that for [50, Proposition 1.2.1], we can prove that the subsequence of final stepsizes $\{\eta_k\}_{k \in \mathcal{K}} \to 0$ (see the inequality before [50, eq. (1.17)]). Since $\{\bar{s}_k\}$ is lower bounded (asymptotically), we must have that $\bar{s}_k > \eta_k, \ \forall \ k \in \mathcal{K}, k \geq \bar{k}$ for large enough $\bar{k}$. Thus the corresponding $i_2 > 0$ for all $k \in \mathcal{K}, k \geq \bar{k}$. By the definition of i_2 in (249), we know that $i_2 - 1$ does not satisfy the reduction requirement; in other words, we have $\tilde{F}(\boldsymbol{x}_k(\eta_k \xi^{-1})) > \tilde{F}(\boldsymbol{x}_{k-1}) - \sigma \eta_k \xi^{-1} \|\nabla \tilde{F}(\boldsymbol{x}_{k-1})\|_F^2$, or equivalently,

$$\begin{aligned} &\tilde{F}(\boldsymbol{x}_{k-1}) - \tilde{F}(\boldsymbol{x}_{k-1} - \eta_k \xi^{-1} \nabla \tilde{F}(\boldsymbol{x}_{k-1}))) \\ &\quad < \sigma \eta_k \xi^{-1} \|\nabla \tilde{F}(\boldsymbol{x}_{k-1})\|_F^2, \quad \forall \ k \in \mathcal{K}, k \geq \bar{k}. \end{aligned}$$

This relation is the same as [50, eq. (1.17)] (except that [50, eq. (1.17)] considers a more general descent direction), and the rest of the proof is also the same as [50] and is omitted here.

For Algorithm 1 with stepsize chosen by the restricted line search rule, since it "gives larger reduction in cost at each iteration" than the restricted Armijo rule, it "inherits the convergence properties" of the restricted Armijo rule (as remarked in the last paragraph of the proof of [50, Proposition 1.2.1]). The rigorous proof is similar to that in the second last paragraph of the proof of [50, Proposition 1.2.1]) and is omitted here.

Algorithm 2 is a two-block BCD method to solve problem (P1). According to [59, Corollary 2], each limit point of the sequence generated by Algorithm 2 is a stationary point of problem (P1).

Algorithm 3 belongs to the class of BSUM methods [55]. According to Proposition D.1, the level set $\mathcal{X}^0 = \{\boldsymbol{x} \mid \tilde{F}(\boldsymbol{x}) \leq \tilde{F}(\boldsymbol{x}_0)\}$ is a subset of the bounded set $K_1 \cap K_2$, thus $\mathcal{X}^0$ is bounded. Moreover, $\mathcal{X}^0$ is a closed set, thus $\mathcal{X}^0$ is compact. It is easy to verify that the objective function of each subproblem in Algorithm 3 is a convex tight upper bound of $\tilde{F}(\boldsymbol{x})$ (more precisely, satisfies [55, Assumption 2]). It is also obvious that the objective function of each subproblem is strongly convex, thus each subproblem of Algorithm 3 has a unique solution. Based on these facts, it follows from [55, Th. 2] that each limit point of the sequence generated by Algorithm 3 is a stationary point.

Algorithm 4 is a SGD method (or more precisely, incremental gradient method) with a specific stepsize rule. According to (243) and (244) in Appendix D.4, Algorithm 4 can be viewed as an approximate gradient descent method with bounded error. By [62, Proposition 1], each limit point of the sequence generated by Algorithm 4 is a stationary point.

E. Proof of Lemma 3.3

We will prove a statement that is stronger than Lemma 3.1: with probability at least $1-1/n^4$, for any $(X,Y) \in K_1 \cap K_2 \cap K(\delta)$ and U, V defined in Table VII, we have

$$\begin{aligned}&\langle \nabla_X \tilde{F}(X,Y), X-U\rangle + \langle \nabla_Y \tilde{F}(X,Y), Y-V\rangle \\ &\quad \geq \frac{p}{4}d^2 + \frac{2\sqrt{\rho}}{\Sigma_{\min}} d\sqrt{G(X,Y)}, \end{aligned} \tag{251}$$

where $d = \|M - XY^T\|_F$.

We have already proved (37a), i.e. with probability at least $1-1/n^4$,

$$\phi_F = \langle \nabla_X F, X-U\rangle + \langle \nabla_Y F, Y-V\rangle \geq \frac{p}{4}d^2.$$

It remains to prove a bound on ϕ_G, which is stronger than the bound $\phi_G \geq 0$. Note that ϕ_F depends on the observed set $\boldsymbol{\Omega}$, thus the bound on ϕ_F holds with high probability; in contrast, ϕ_G does not depend on $\boldsymbol{\Omega}$, thus the bound on ϕ_G always holds.

Claim E.1: For any $(X,Y) \in K_1 \cap K_2 \cap K(\delta)$ and U, V defined in Table VII, we have

$$\phi_G = \langle \nabla_X G, X-U\rangle + \langle \nabla_Y G, Y-V\rangle \geq \frac{2\sqrt{\rho}}{\Sigma_{\min}} d\sqrt{G(X,Y)}. \tag{252}$$

Proof of Claim E.1: By the definition of G in (13), $G(X,Y) = \rho(\sum_i G_{1i}(X) + G_2(X) + \sum_j G_{3j}(Y) + G_4(Y))$, where the component functions

$$\begin{aligned} G_{1i}(X) &= G_0\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right), \quad G_2(X) = G_0\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right), \\ G_{3j}(Y) &\triangleq G_0\left(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2}\right), \quad G_4(Y) \triangleq G_0\left(\frac{3\|Y\|_F^2}{2\beta_T^2}\right). \end{aligned} \tag{253}$$

By the expressions of $\nabla_X G, \nabla_Y G$ in (24), we have

$$\begin{aligned} \phi_G &= \langle \nabla_X G, X-U\rangle + \langle \nabla_Y G, Y-V\rangle \\ &= \rho \sum_{i=1}^m G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2})\frac{3}{\beta_1^2}\langle X^{(i)}, X^{(i)} - U^{(i)}\rangle \\ &\quad + \rho G_0'(\frac{3\|X\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle X, X-U\rangle \\ &\quad + \rho \sum_{j=1}^n G_0'(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2})\frac{3}{\beta_2^2}\langle Y^{(j)}, Y^{(j)} - V^{(j)}\rangle \\ &\quad + \rho G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle Y, Y-V\rangle, \end{aligned} \tag{254}$$

where $G_0'(z) = I_{[1,\infty]}(z)2(z-1) = 2\sqrt{G_0(z)}$.

Firstly, we prove

$$\begin{aligned} h_{1i} &\triangleq G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2})\frac{3}{\beta_1^2}\langle X^{(i)}, X^{(i)} - U^{(i)}\rangle \\ &\geq \frac{1}{2}\sqrt{G_{1i}(X)}, \quad \forall\ i, \end{aligned} \tag{255a}$$

$$\begin{aligned} h_{3j} &\triangleq G_0'(\frac{3\|Y^{(j)}\|^2}{2\beta_2^2})\frac{3}{\beta_2^2}\langle Y^{(j)}, Y^{(j)} - V^{(j)}\rangle \\ &\geq \frac{1}{2}\sqrt{G_{3j}(Y)}, \quad \forall\ j. \end{aligned} \tag{255b}$$

We only need to prove (255a); the proof of (255b) is similar. We consider two cases.

Case 1: $\|X^{(i)}\|^2 \leq \frac{2\beta_1^2}{3}$. Note that $\frac{3\|X^{(i)}\|^2}{2\beta_1^2} \leq 1$ implies $G_0(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}) = G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}) = 0$, thus $h_{1i} = G_{1i} = 0$, in which case (255a) holds.

Case 2: $\|X^{(i)}\|^2 > \frac{2\beta_1^2}{3}$. By Corollary 4.1 and the fact that $\beta_1^2 = \beta_T^2 \frac{3\mu r}{m}$, we have

$$\|U^{(i)}\|^2 \leq \frac{3r\mu}{2m}\beta_T^2 \overset{(15)}{=} \frac{3}{4}\frac{2\beta_1^2}{3} < \frac{3}{4}\|X^{(i)}\|^2. \tag{256}$$

As a result, $\frac{\sqrt{3}}{2}\langle X^{(i)}, X^{(i)}\rangle = \frac{\sqrt{3}}{2}\|X^{(i)}\|\|X^{(i)}\| > \|X^{(i)}\|\|U^{(i)}\| \geq \langle X^{(i)}, U^{(i)}\rangle$, which implies $\langle X^{(i)}, X^{(i)} - U^{(i)}\rangle \geq (1-\frac{\sqrt{3}}{2})\|X^{(i)}\|^2 > (1-\frac{\sqrt{3}}{2})\frac{2}{3}\beta_1^2 > \frac{1}{12}\beta_1^2$. Combining this inequality with the fact that $G_0'(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}) = 2\sqrt{G_0\left(\frac{3\|X^{(i)}\|^2}{2\beta_1^2}\right)} = 2\sqrt{G_{1i}(X)}$, we get (255a).

Secondly, we prove

$$h_2 + h_4 \geq \frac{2d}{\Sigma_{\min}} \left(\sqrt{G_2(X)} + \sqrt{G_4(Y)}\right),$$
$$\text{where } \; h_2 \triangleq G_0'(\frac{3\|X\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle X, X-U\rangle,$$
$$h_4 \triangleq G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle Y, Y-V\rangle. \tag{257}$$

Without loss of generality, we can assume $\|Y\|_F \geq \|X\|_F$, and we will apply Corollary 4.1 to prove (257). If $\|Y\|_F < \|X\|_F$, we can apply a symmetric result of Corollary 4.1 to prove (257). We consider three cases.

Case 1: $\|X\|_F \leq \|Y\|_F \leq \sqrt{\frac{2}{3}}\beta_T$. In this case $G_0(\frac{3\|X\|_F^2}{2\beta_T^2}) = G_0'(\frac{3\|X\|_F^2}{2\beta_T^2}) = G_0(\frac{3\|Y\|_F^2}{2\beta_T^2}) = G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2}) = 0$, which implies $h_2 = h_4 = G_2(X) = G_4(Y) = 0$, thus (257) holds.

Case 2: $\|X\|_F \leq \sqrt{\frac{2}{3}}\beta_T < \|Y\|_F$. Then we have $\frac{3\|X\|_F^2}{2\beta_T^2} \leq 1$, which implies $h_2 = 0 = G_2(X)$. By (51d) in Corollary 4.1 we have $\|V\|_F \leq (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F$, which implies $(1 - \frac{d}{\Sigma_{\min}})\langle Y, Y\rangle = (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F^2 \geq \|Y\|_F\|V\|_F \geq \langle Y, V\rangle$. This further implies $\langle Y, Y - V\rangle \geq \frac{d}{\Sigma_{\min}}\|Y\|_F^2 \geq \frac{d}{\Sigma_{\min}}\frac{2\beta_T^2}{3}$. Combined with the fact that $G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2}) = 2\sqrt{G_0(\frac{3\|Y\|_F^2}{2\beta_T^2})} = 2\sqrt{G_4(Y)}$, we get

$$h_4 = G_0'(\frac{3\|Y\|_F^2}{2\beta_T^2})\frac{3}{\beta_T^2}\langle Y, Y-V\rangle$$
$$\geq 2\sqrt{G_4(Y)}\frac{3}{\beta_T^2}\frac{d}{\Sigma_{\min}}\frac{2\beta_T^2}{3} = \frac{4d}{\Sigma_{\min}}\sqrt{G_4(Y)}.$$

Thus $h_2 + h_4 = h_4 \geq \frac{4d}{\Sigma_{\min}}\sqrt{G_4(Y)} = \frac{4d}{\Sigma_{\min}}\left(\sqrt{G_4(Y)} + \sqrt{G_2(X)}\right) \geq \frac{2d}{\Sigma_{\min}}\left(\sqrt{G_4(Y)} + \sqrt{G_2(X)}\right)$.

Case 3: $\sqrt{\frac{2}{3}}\beta_T < \|X\|_F \leq \|Y\|_F$. Since $\|Y\|_F \geq \|X\|_F$, we have $G_4(Y) = G_0\left(\frac{3\|Y\|_F^2}{2\beta_T^2}\right) \geq G_0\left(\frac{3\|X\|_F^2}{2\beta_T^2}\right) = G_2(X)$. By Corollary 4.1, we have $\|U\|_F \leq \|X\|_F$ and $\|V\|_F \leq (1 - \frac{d}{\Sigma_{\min}})\|Y\|_F$. Similar to the argument in Case 2 we can prove $h_2 \geq 0, h_4 \geq \frac{4d}{\Sigma_{\min}}\sqrt{G_4(Y)}$; thus $h_2 + h_4 \geq \frac{4d}{\Sigma_{\min}}\sqrt{G_4(Y)} \geq \frac{2d}{\Sigma_{\min}}\left(\sqrt{G_4(Y)} + \sqrt{G_2(X)}\right)$.

In all three cases, we have proved (257), thus (257) holds.

We conclude that for U, V defined in Table VII, (258), as shown at the bottom of this page, which finishes the proof of Claim E.1. □

Let us come back to the proof of Lemma 3.3. The rest of the proof is just algebraic computation. According to (251), we have

$$\begin{aligned}
&\frac{p}{4}d^2 + \frac{2\sqrt{\rho}}{\Sigma_{\min}}d\sqrt{G(X,Y)} \\
&\leq \langle \nabla_X \tilde{F}(X,Y), X-U\rangle + \langle \nabla_Y \tilde{F}(X,Y), Y-V\rangle \\
&\leq (\|\nabla_X \tilde{F}(X,Y)\|_F + \|\nabla_Y \tilde{F}(X,Y)\|_F) \\
&\quad \max\{\|X-U\|_F, \|Y-V\|_F\} \\
&\overset{(51b)}{\leq} \sqrt{2}\sqrt{\|\nabla_X \tilde{F}(X,Y)\|_F^2 + \|\nabla_Y \tilde{F}(X,Y)\|_F^2}\frac{17}{2}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d \\
&= \|\nabla \tilde{F}(X,Y)\|_F \frac{17}{\sqrt{2}}\sqrt{r}\frac{\beta_T}{\Sigma_{\min}}d.
\end{aligned}$$

Eliminating a factor of d from both sides and taking square, we get

$$\begin{aligned}
\|\nabla \tilde{F}(X,Y)\|_F^2 \frac{289}{2} r \frac{\beta_T^2}{\Sigma_{\min}^2} &\geq \left(\frac{p}{4}d + \frac{2\sqrt{\rho}}{\Sigma_{\min}}\sqrt{G(X,Y)}\right)^2 \\
&\geq \frac{pd^2}{16} + \frac{4\rho}{\Sigma_{\min}^2}G(X,Y).
\end{aligned} \tag{259}$$

By the definition of β_T in (15), we have

$$r\frac{\beta_T^2}{\Sigma_{\min}^2} = r\frac{C_T r \Sigma_{\max}}{\Sigma_{\min}^2} = C_T \frac{r^2 \kappa}{\Sigma_{\min}}.$$

According to Claim 3.1, we have

$$pd^2 = p\|M - XY^T\|_F^2 \geq \frac{1}{2}\|\mathcal{P}_\Omega(M - XY^T)\|_F^2 = F(X,Y).$$

$$\begin{aligned}
\phi_G &\overset{(254)}{=} \rho\left(\sum_i h_{1i} + \sum_j h_{3j} + h_2 + h_4\right) \\
&\overset{(255),(257)}{\geq} \rho\left(\frac{1}{2}\sum_i \sqrt{G_{1i}(X)} + \frac{1}{2}\sum_j \sqrt{G_{2j}(Y)}\right. \\
&\quad \left. + \frac{2d}{\Sigma_{\min}}\sqrt{G_2(X)} + \frac{2d}{\Sigma_{\min}}\sqrt{G_4(Y)}\right) \\
&\geq \rho\frac{2d}{\Sigma_{\min}}\left(\sum_i \sqrt{G_{1i}(X)} + \sum_j \sqrt{G_{2j}(Y)} + \sqrt{G_2(X)} + \sqrt{G_4(Y)}\right) \\
&\geq \rho\frac{2d}{\Sigma_{\min}}\sqrt{\sum_i G_{1i}(X) + \sum_j G_{2j}(Y) + G_2(X) + G_4(Y)} \\
&= \rho\frac{2d}{\Sigma_{\min}}\sqrt{\frac{1}{\rho}G(X,Y)} = \frac{2\sqrt{\rho}}{\Sigma_{\min}}d\sqrt{G(X,Y)}.
\end{aligned} \tag{258}$$

By the definition of ρ in (17) and the definition of δ_0 in (16), we have

$$\frac{4\rho}{\Sigma_{\min}^2} = \frac{4}{\Sigma_{\min}^2} 8p\delta_0^2 = \frac{32p}{\Sigma_{\min}^2} \frac{1}{36} \frac{\Sigma_{\min}^2}{C_d^2 r^3 \kappa^2} = \frac{8}{9} \frac{1}{C_d^2 r^3 \kappa^2} p.$$

Substituting the above three relations into (259), we get (when $C_d \geq 32/3$)

$$\begin{aligned}
&\|\nabla \tilde{F}(X,Y)\|_F^2 \frac{289}{2} C_T \frac{r^2\kappa}{\Sigma_{\min}} \\
&\geq \frac{p}{32} F(X,Y) + \frac{8}{9} \frac{1}{C_d^2 r^3 \kappa^2} pG(X,Y) \\
&\geq \frac{8}{9} \frac{1}{C_d^2 r^3 \kappa^2} p(F(X,Y) + G(X,Y)) = \frac{8}{9} \frac{1}{C_d^2 r^3 \kappa^2} p\tilde{F}(X,Y).
\end{aligned}$$

This can be further simplified to

$$\|\nabla \tilde{F}(X,Y)\|_F^2 \geq \frac{\Sigma_{\min}}{C_g r^5 \kappa^3} p\tilde{F}(X,Y),$$

where the numerical constant $C_g = \frac{2601}{16} C_T C_d^2$. This finishes the proof of Lemma 3.3.

REFERENCES

[1] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *IEEE Comput.*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[2] P. Chen and D. Suter, "Recovering the missing components in a large noisy low-rank matrix: Application to SFM," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1051–1063, Aug. 2004.

[3] Z. Liu and L. Vandenberghe, "Interior-point method for nuclear norm approximation with application to system identification," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 3, pp. 1235–1256, 2009.

[4] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[5] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.

[6] D. Gross, "Recovering low-rank matrices from few coefficients in any basis," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1548–1566, Mar. 2011.

[7] B. Recht, "A simpler approach to matrix completion," *J. Mach. Learn. Res.*, vol. 12, pp. 3413–3430, Jan. 2011.

[8] E. J. Candès and Y. Plan, "Matrix completion with noise," *Proc. IEEE*, vol. 98, no. 6, pp. 925–936, Jun. 2010.

[9] S. Negahban and M. J. Wainwright, "Restricted strong convexity and weighted matrix completion: Optimal bounds with noise," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 1665–1697, 2012.

[10] J.-F. Cai, E. J. Candès, Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.

[11] S. Ma, D. Goldfarb, and L. Chen, "Fixed point and Bregman iterative methods for matrix rank minimization," *Math. Program.*, vol. 128, nos. 1–2, pp. 321–353, 2011.

[12] K. C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific J. Optim.*, vol. 6, nos. 615–640, p. 15, 2010.

[13] A. Agarwal, S. Negahban, and M. J. Wainwright, "Fast global convergence of gradient methods for high-dimensional statistical recovery," *Ann. Statist.*, vol. 40, no. 5, pp. 2452–2482, 2012.

[14] K. Hou, Z. Zhou, A. M.-C. So, and Z.-Q. Luo, "On the linear convergence of the proximal gradient method for trace norm regularization," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2013, pp. 710–718.

[15] A. P. Singh and G. J. Gordon, "A unified view of matrix factorization models," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*, 2008, pp. 358–373.

[16] G. Takács, I. Pilászy, B. Németh, and D. Tikk, "Major components of the gravity recommendation system," *ACM SIGKDD Explorations Newslett.*, vol. 9, no. 2, pp. 80–83, 2007.

[17] R. H. Keshavan, "Efficient algorithms for collaborative filtering," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, 2012.

[18] P. Jain, P. Netrapalli, and S. Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proc. 45th Annu. ACM Symp. Theory Comput. (STOC)*, 2013, pp. 665–674.

[19] M. Hardt, "Understanding alternating minimization for matrix completion," in *Proc. IEEE 55th Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2014, pp. 651–660.

[20] M. Hardt and M. Wootters, "Fast matrix completion without the condition number," in *Proc. 27th Conf. Learn. Theory (COLT)*, 2014, pp. 638–678.

[21] Y. Zhou, D. Wilkinson, R. Schreiber, and R. Pan, "Large-scale parallel collaborative filtering for the Netflix prize," in *Proc. Int. Conf. Algorithmic Appl. Manage.*, 2008, pp. 337–348.

[22] Z. Wen, W. Yin, and Y. Zhang, "Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm," *Math. Program. Comput.*, vol. 4, no. 4, pp. 333–361, 2012.

[23] S. Funk. *Netflix Update: Try This at Home*, accessed on Dec. 2006. [Online]. Available: http://sifter.org/~simon/journal/20061211.html

[24] A. Paterek, "Improving regularized singular value decomposition for collaborative filtering," in *Proc. KDD Cup Workshop*, vol. 2007. 2007, pp. 5–8.

[25] R. Gemulla, E. Nijkamp, P. J. Haas, and Y. Sismanis, "Large-scale matrix factorization with distributed stochastic gradient descent," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 69–77.

[26] B. Recht and C. Ré, "Parallel stochastic gradient algorithms for large-scale matrix completion," *Math. Program. Comput.*, vol. 5, no. 2, pp. 201–226, 2013.

[27] Y. Zhuang, W.-S. Chin, Y.-C. Juan, and C.-J. Lin, "A fast parallel SGD for matrix factorization in shared memory systems," in *Proc. 7th ACM Conf. Recommender Syst.*, 2013, pp. 249–256.

[28] I. Pilászy, D. Zibriczky, and D. Tikk, "Fast ALS-based matrix factorization for explicit and implicit feedback datasets," in *Proc. 4th ACM Conf. Recommender Syst.*, 2010, pp. 71–78.

[29] H.-F. Yu, C.-J. Hsieh, S. Si, and I. Dhillon, "Scalable coordinate descent approaches to parallel matrix factorization for recommender systems," in *Proc. ICDM*, 2012, pp. 765–774.

[30] R. Sun, "Matrix completion via nonconvex factorization: Algorithms and theory," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Minnesota, Minneapolis, MN, USA, 2015.

[31] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2980–2998, Jun. 2010.

[32] E. J. Candès, X. Li, and M. Soltanolkotabi. (2014). "Phase retrieval via Wirtinger flow: Theory and algorithms." [Online]. Available: https://arxiv.org/abs/1407.1065

[33] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert. (2009). "Quantum state tomography via compressed sensing." [Online]. Available: http://arxiv.org/abs/0909.3304v1

[34] P. Jain and P. Netrapalli. (2014). "Fast exact matrix completion with finite samples." [Online]. Available: http://arxiv.org/abs/1411.1087

[35] C. De Sa, K. Olukotun, and C. Ré. (2014). "Global convergence of stochastic gradient descent for some non-convex matrix problems." [Online]. Available: https://arxiv.org/abs/1411.1134

[36] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2013, pp. 2796–2804.

[37] C.-H. Zhang and T. Zhang, "A general theory of concave regularization for high-dimensional sparse estimation problems," *Statist. Sci.*, vol. 27, no. 4, pp. 576–593, 2012.

[38] P.-L. Loh and M. Wainwright, "Regularized M-estimators with nonconvexity: Statistical and algorithmic theory for local optima," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 476–484.

[39] J. Fan, L. Xue, and H. Zou, "Strong oracle optimality of folded concave penalized estimation," *Ann. Statist.*, vol. 42, no. 3, pp. 819–849, 2014.

[40] X.-T. Yuan and T. Zhang, "Truncated power method for sparse eigenvalue problems," *J. Mach. Learn. Res.*, vol. 14, no. 1, pp. 899–925, 2013.

[41] Z. Wang, H. Lu, and H. Liu. (2014). "Nonconvex statistical optimization: Minimax-optimal sparse PCA in polynomial time." [Online]. Available: https://arxiv.org/abs/1408.5352

[42] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain, "Non-convex robust PCA," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1107–1115.

[43] S. Balakrishnan, M. J. Wainwright, and B. Yu. (2014). "Statistical guarantees for the EM algorithm: From population to sample-based analysis." [Online]. Available: https://arxiv.org/abs/1408.2156
[44] Z. Wang, Q. Gu, Y. Ning, and H. Liu. (2014). "High dimensional expectation-maximization algorithm: Statistical optimization and asymptotic normality." [Online]. Available: https://arxiv.org/abs/1412.8729
[45] P.-Å. Wedin, "Perturbation bounds in connection with singular value decomposition," *BIT Numer. Math.*, vol. 12, no. 1, pp. 99–111, 1972.
[46] U. Feige and E. Ofek, "Spectral techniques applied to sparse random graphs," *Random Struct. Algorithms*, vol. 27, no. 2, pp. 251–275, Sep. 2005.
[47] Y. Chen, S. Bhojanapalli, S. Sanghavi, and R. Ward, "Coherent matrix completion," in *Proc. 31st Int. Conf. Mach. Learn. (ICML)*, 2014, pp. 674–682.
[48] S. Bhojanapalli and P. Jain. (2014). "Universal matrix completion." [Online]. Available: http://arxiv.org/abs/1402.2324
[49] W. I. Zangwill, "Non-linear programming via penalty functions," *Manage. Sci.*, vol. 13, no. 5, pp. 344–358, 1967.
[50] D. P. Bertsekas, *Nonlinear Programming*. Belmont, CA, USA: Athena Scientific Belmont, 1999.
[51] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, no. 3, pp. 475–494, 2001.
[52] R. Sun and M. Hong, "Improved iteration complexity bounds of cyclic block coordinate descent for convex problems," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1306–1314.
[53] R. Sun and Y. Ye. (2016). "Worst-case complexity of cyclic coordinate descent: $O(n^2)$ gap with randomized version." [Online]. Available: https://arxiv.org/abs/1604.07130
[54] Y. Nesterov, "Efficiency of coordinate descent methods on huge-scale optimization problems," *SIAM J. Optim.*, vol. 22, no. 2, pp. 341–362, 2012.
[55] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.
[56] H. Baligh *et al.*, "Cross-layer provision of future cellular networks: A WMMSE-based approach," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 56–68, Nov. 2014.
[57] M. Hong, R. Sun, H. Baligh, and Z. Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 226–240, Feb. 2013.
[58] T. Hastie, R. Mazumder, J. Lee, and R. Zadeh. (2014). "Matrix completion and low-rank SVD via fast alternating least squares." [Online]. Available: https://arxiv.org/abs/1410.2596
[59] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss–Seidel method under convex constraints," *Oper. Res. Lett.*, vol. 26, no. 3, pp. 127–136, Apr. 2000.
[60] R. Sun, Z.-Q. Luo, and Y. Ye. (2015). "On the expected convergence of randomly permuted ADMM." [Online]. Available: https://arxiv.org/abs/1503.06387
[61] L. Zhi-Quan and T. Paul, "Analysis of an approximate gradient projection method with applications to the backpropagation algorithm," *Optim. Methods Softw.*, vol. 4, no. 2, pp. 85–101, 1994.
[62] D. P. Bertsekas and J. N. Tsitsiklis, "Gradient convergence in gradient methods with errors," *SIAM J. Optim.*, vol. 10, no. 3, pp. 627–642, 2000.
[63] G. W. Stewart, *Perturbation Theory for the Singular Value Decomposition*. 1998.

Ruoyu Sun is currently a visiting researcher at FAIR (Facebook Artificial Intelligence Research). He will join UIUC IS&E department as an assistant professor in 2017. Previously he was a postdoctoral scholar in the Department of Management Science & Engineering at Stanford University. He obtained his Ph.D. in Electrical Engineering at the University of Minnesota in 2015, under the supervision of Zhi- Quan (Tom) Luo. He received the B.Sc. degree in mathematics from Peking University, Beijing, China in 2009.

His research interest lies in large-scale optimization, machine learning, signal processing, information theory, wireless communications. He received the second place of 2015 INFORMS Nicholson student paper competition, and the honorable mention of 2015 INFORMS optimization society student paper prize for this paper.

Zhi-Quan Luo (M'90–SM'02–F'07) received his B.Sc. degree in Applied Mathematics in 1984 from Peking University, Beijing, China. Subsequently, he was selected by a joint committee of the American Mathematical Society and the Society of Industrial and Applied Mathematics to pursue Ph.D study in the United States. After an one-year intensive training in mathematics and English at the Nankai Institute of Mathematics, Tianjin, China, he studied in the Operations Research Center and the Department of Electrical Engineering and Computer Science at MIT, where he received a Ph.D degree in Operations Research in 1989.

From 1989 to 2003, Dr. Luo held a faculty position with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada, where he eventually became the department head and held a Canada Research Chair in Information Processing. Since April of 2003, he has been with the Department of Electrical and Computer Engineering at the University of Minnesota (Twin Cities) as a full professor. His research interests lies in the union of optimization algorithms, signal processing and digital communication. Currently, he is with the Chinese University of Hong Kong, Shenzhen, where he is a Professor and serves as the Vice President Academic.

Dr. Luo is a fellow of IEEE and SIAM. He is a recipient of the 2004, 2009 and 2011 IEEE Signal Processing Society's Best Paper Awards, the 2011 EURASIP Best Paper Award and the 2011 ICC Best Paper Award. He was awarded the 2010 Farkas Prize from the INFORMS Optimization Society. Dr. Luo chaired the IEEE Signal Processing Society Technical Committee on the Signal Processing for Communications (SPCOM) from 2011-2012. He has held editorial positions for several international journals including *Journal of Optimization Theory and Applications*, *SIAM Journal on Optimization*, *Mathematics of Computation* and IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is the current Editor-in-Chief for the journal IEEE Trans. Signal Processing. In 2014 he is elected to the Royal Society of Canada.